Psychometrika

CONTENTS

DETERMINATION OF OPTIMAL TEST LENGTH TO MAXIMIZE THE MULTIPLE CORRELATION - - - 79
PAUL HORST

NOTE ON THE COMPUTATION OF THE INVERSE OF A TRIANGULAR MATRIX - - - - - - - - 89
BENJAMIN FRUCHTER

A METHOD OF MATRIX ANALYSIS OF GROUP STRUCTURE - - - - - - - - - - - - - - - - 95
R. DUNCAN LUCE AND ALBERT D. PERRY

A NOTE ON THE ESTIMATION OF TEST RELIABILITY BY THE KUDER-RICHARDSON FORMULA (20) - 117
LEDYARD R TUCKER

APPLICATION OF THE CONCEPT OF SIMPLE STRUCTURE TO ALEXANDER'S DATA - - - - - - 121
MARIANO YELA

DEVELOPMENT OF A METHOD FOR INCREASING THE UTILITY OF MULTIPLE CORRELATIONS BY CONSIDERING BOTH TESTING TIME AND TEST VALIDITY - - - - - - - - - - - - - - 137
W. F. LONG AND IRVING W. BURR

DON LEWIS. *Quantitative Methods in Psychology.*
A Review - - - - - - - - - - - - - - - 163
LLOYD G. HUMPHREYS AND LYLE V. JONES

EGON BRUNSWIK. *Systematic and representative design of psychological experiments.*
A Review - - - - - - - - - - - - - - - 165
LEO POSTMAN

VOLUME FOURTEEN JUNE 1949 NUMBER TWO

PSYCHOMETRIKA—VOL. 14, NO. 2
JUNE, 1949

DETERMINATION OF OPTIMAL TEST LENGTH TO MAXIMIZE THE MULTIPLE CORRELATION

PAUL HORST
UNIVERSITY OF WASHINGTON

If the lengths of the tests in a battery are altered, their intercorrelations and their validities or correlations with a criterion are also altered. Consequently, the multiple correlation of the battery with the criterion will also be altered. These changes are a function of the reliabilities of the tests. Suppose we have given from a set of experimental data (1) the time allowed for each test in the battery, (2) the reliability of each test, (3) the intercorrelations, and (4) the validities of all the tests. If we specify the over-all testing time we are willing to allow for the test in the future, we can determine the amount by which each test must be altered in order to give the maximum multiple correlation with the criterion. The method is presented, together with numerical examples and the mathematical proof.

I. *The Method*

In general, when we prepare a battery of tests to predict a given criterion the resulting multiple correlation is not the highest we could get for a given amount of testing time. By altering the lengths of the tests we change their reliabilities and consequently also their validities and intercorrelations. Alteration of the intercorrelations and validities will also change the multiple correlation. Since the amount of administration time required of a test battery is often of considerable importance, it would be well to know for any particular battery whether we could readjust the lengths of the tests without increasing the over-all testing time so as to increase the multiple correlation with the criterion. Suppose we have given the following data:

1. The time required for each test
2. The reliability of each test
3. The intercorrelations of each test with every other
4. The validity of each test
5. The total time we are willing to assign to the battery. This may or may not be the same as the time assigned to the original battery.

With these data we can determine how to alter the time for each test in order to maximize the multiple correlation of the altered battery with the criterion. We assume in this development that if the length of a test is altered by a given percentage, the time allotted to the test is altered by the same percentage so that for our purposes "length of test" and "time for test" are regarded as interchangeable. We assume also, of course, that in altering the time for each test we do not change the function which it measures. Specifically, we assume that in changing the length of a test its correlation with another test will be indicated by

$$R_{12} = \sqrt{\frac{a}{1 + (a-1) r_{11}}}\, r_{12},$$

where

r_{12} is the original correlation between the two tests,

a is the new length of the first test divided by its original length,

r_{11} is the reliability of the first test before its length is altered,

R_{12} is the correlation between the two tests aftr the first has been altered.

If both tests have been altered, then the new correlation will be

$$R_{12} = \sqrt{\frac{ab}{[1 + (a-1) r_{11}][1 + (b-1) r_{22}]}}\, r_{12}$$

where now b is the proportional change in the second test and r_{22} its original reliability.

Both of these formulas are well known and can be found in any good text in psychological or educational statistics.

First we shall indicate the procedure for determining the altered lengths of the tests, including a numerical example. Later the proof of the method will be developed. We let

a_i be the proportion of the original total testing time required by test i,

r_{ii} be the reliability of test i,

r_{ij} be the correlation of test i with test j,

r_{ic} be the validity of test i or its correlation with the criterion,

T be the ratio of the new total testing time to the old,

b_i be T times the proportion of the new testing time required by test i.

From these definitions it should be clear that

$$\sum a = 1$$
$$\sum b = T\,.$$

Furthermore, we let

$$u_i = 1 - r_{ii}$$

and call u_i the unreliability of test i.

We shall first indicate the procedure for determining the b_i's in terms of a solution for three independent variables. The solution can be generalized to any number of variables. First we write the equations

$$\left.\begin{array}{l}
\left(r_{11}+\dfrac{a_1u_1}{T}\right)\beta_1+\left(r_{12}+\dfrac{\sqrt{a_1u_1a_2u_2}}{T}\right)\beta_2+\left(r_{13}+\dfrac{\sqrt{a_1u_1a_3u_3}}{T}\right)\beta_3=r_{1c}\\[2ex]
\left(r_{12}+\dfrac{\sqrt{a_1u_1a_2u_2}}{T}\right)\beta_1+\left(r_{22}+\dfrac{a_2u_2}{T}\right)\beta_2+\left(r_{23}+\dfrac{\sqrt{a_2u_2a_3u_3}}{T}\right)\beta_3=r_{2c}\\[2ex]
\left(r_{13}+\dfrac{\sqrt{a_1u_1a_3u_3}}{T}\right)\beta_1+\left(r_{23}+\dfrac{\sqrt{a_2u_2a_3u_3}}{T}\right)\beta_2+\left(r_{33}=\dfrac{a_3u_3}{T}\right)\beta_3=r_{3c}
\end{array}\right\}. \quad (1)$$

It will be noted that equations (1) are somewhat similar to a set of normal equations used in solving for conventional β regression weights. The right-hand sides of the equations are the validity coefficients. The left-hand sides include the intercorrelations. However, the diagonal terms on the left contain the reliabilities instead of unity. Furthermore, to each coefficient is added a term of the type $\dfrac{\sqrt{a_iu_ia_ju_j}}{T}$. It will be noted that in the diagonal terms, $i = j$, and therefore $\dfrac{\sqrt{a_iu_ia_ju_j}}{T}$ is simply $\dfrac{a_iu_i}{T}$.

If we assume now that all the values in equation (1) except the β's are known, we can solve for the β's by any method desired. After solving for the β's we solve for the b's by means of the equations

$$\left.\begin{aligned} b_1 &= \beta_1 \sqrt{u_1 a_1}\,\frac{T}{\sum_1^3 \beta_i \sqrt{u_i a_i}} \\ b_2 &= \beta_2 \sqrt{u_2 a_2}\,\frac{T}{\sum_1^3 \beta_i \sqrt{u_i a_i}} \\ b_3 &= \beta_3 \sqrt{u_3 a_3}\,\frac{T}{\sum_1^3 \beta_i \sqrt{u_i a_i}} \end{aligned}\right\} . \tag{2}$$

The b's in equations (2) tell us now by what proportion to change the time of each test in order that the total testing time be T times that of the original, and the multiple correlation of the tests with the criterion be a maximum.

The actual value of the new multiple correlation can be determined by the equation

$$R_b^2 = \beta_1 r_{1c} + \beta_2 r_{2c} + \beta_3 r_{3c} . \tag{3}$$

Formula (3) is analogous to many published formulas.

We shall now illustrate the procedure numerically with an example having two independent variables. We let

$$\begin{array}{ll} a_1 = .2 & a_2 = .8 \\ r_{11} = .60 & r_{22} = .80 \\ r_{12} = .20 & \\ r_{1c} = .40 & r_{2c} = .30 . \end{array}$$

Then

$$\begin{array}{ll} u_1 = .40 & u_2 = .20 . \end{array}$$

First we shall take the case of $T = 1$; that is, we do not wish to change the over-all time for the test. Our equations (1) would then be

$$(.60 + .4 \times .2)\beta_1 + (.20 + \sqrt{.4 \times .2 \times .2 \times .8})\beta_2 = .40$$

$$(.20 + \sqrt{.4 \times .2 \times .2 \times .8})\beta_2 + (.80 + .2 \times .8)\beta_2 = .30$$

or

$$.680\,\beta_1 + .313\,\beta_2 = .40$$

$$.313\,\beta_1 + .96\ \ \beta_2 = .30 .$$

Solving these equations for β_1 and β_2 we have

$$\beta_1 = .523\,, \qquad \beta_2 = .142\,.$$

Using equations (2) to solve for the b's,

$$b_1 = \frac{.523 \times \sqrt{.08}}{.523 \times \sqrt{.08} + .142\,\sqrt{.16}} = .722\,,$$

$$b_2 = \frac{.142 \times \sqrt{.16}}{.523 \times \sqrt{.08} + .142\,\sqrt{.16}} = .278.$$

We see then that the relative administration times of .2 and .8 for tests 1 and 2, respectively, should be changed to .722 and .278.

To calculate the ratio of the new times to the old, we have

$$\frac{b_1}{a_1} = \frac{.722}{.2} = 3.61\,,$$

$$\frac{b_2}{a_2} = \frac{.278}{.8} = \ .35.$$

Therefore test 1 would be more than tripled, while test 2 would be only about one-third of its original length.

The square of the multiple correlation for the altered tests would be as given by equation 3.

$$R_b{}^2 = .523 \times .4 + .142 \times .3 = .25\,.$$

This compares with .21 for the tests of original length and represents an improvement of about 20% .

Suppose now we are willing to double the testing time so that $T = 2$. Using the same numerical example, we have for equations (1)

$$\left(.60 + \frac{.4 \times .2}{2}\right)\beta_1 + \left(.20 + \frac{\sqrt{.4 \times .2 \times .2 \times .8}}{2}\right)\beta_2 = .40$$

$$\left(.20 + \frac{\sqrt{.4 \times .2 \times .2 \times .8}}{2}\right)\beta_1 + \left(.80 + \frac{.2 \times .8}{2}\right)\beta_2 = .30$$

or

$$\begin{aligned} .640\,\beta_1 + .256\,\beta_2 &= .40 \\ .256\,\beta_1 + .880\,\beta_2 &= .30\,. \end{aligned}$$

Solving these equations for the β's, we have

$$\beta_1 = .552\,, \qquad \beta_2 = .181\,.$$

Using equations (2) to solve for the b's,

$$b_1 = \frac{.552 \times \sqrt{.08} \times 2}{.552 \times \sqrt{.08} + .181 \times \sqrt{.16}} = 1.368,$$

$$b_2 = \frac{.181 \times \sqrt{.16} \times 2}{.552 \times \sqrt{.08} + .181 \times \sqrt{.16}} = .632.$$

We have then

$$\frac{b_1}{a_1} = \frac{1.368}{.2} = 6.84,$$

$$\frac{b_2}{a_2} = \frac{.632}{.8} = .79.$$

Therefore, if we double the original testing time, test 1 would be increased almost seven times while test 2 would be reduced to only 79% of its original time.

If we double the over-all testing time, we have for the square of the multiple correlation

$$R_b^2 = .552 \times .4 + .181 \times .3 = .275$$

as compared with the original of .21.

It should be pointed out that the actual computations involved in the procedure for determining optimal test length are negligibly greater than when regression weights and the multiple correlation are calculated from the original data. Furthermore, these computations yield also the regression weights and the multiple correlation coefficients for the tests of altered length.

II. *Proof of the Method*

We let

$r =$ the matrix of intercorrelations of the original tests,

$\rho =$ the matrix of intercorrelations of the altered tests,

$r_c =$ the vector of validity coefficients of the original tests,

$\rho_c =$ the vector of validity coefficients of altered tests,

$D_a =$ a diagonal matrix of the lengths of the original tests,

$D_b =$ a diagonal matrix of the lengths of altered tests,

$D_e = D_b D_a^{-1} =$ a diagonal matrix of the ratios of the new to the old lengths,

$D_{r_{ii}} =$ a diagonal matrix of the test reliabilities.

We let

$$\delta = [I + (D_e - I) D_{r_{ii}}] D_e^{-1}. \tag{1}$$

It can readily be proved then that

$$\rho_c = \delta^{-\frac{1}{2}} r_c , \tag{2}$$

for the typical element of (2) is

$$\rho_{ci} = \sqrt{\frac{e_i}{1 + (e_i - 1) r_{ii}}}\, r_{ci} , \tag{3}$$

which is a well-known formula.

Similarly, it can also be proved that

$$\rho = \delta^{-\frac{1}{2}} (r + d) \delta^{-\frac{1}{2}} , \tag{4}$$

where d is a diagonal matrix which will make the diagonals of ρ unity, that is,

$$\delta^{-\frac{1}{2}} (I + d) \delta^{-\frac{1}{2}} = I \tag{5}$$

or

$$d = \delta - I . \tag{6}$$

From (1) and (6)

$$d = [(I - D_{r_{ii}}) + D_e (D_{r_{ii}} - I)] D_e^{-1} . \tag{7}$$

Let

$$D_u = I - D_{r_{ii}} \tag{8}$$

be a diagonal matrix of the unreliabilities of the tests. From (7) and (8)

$$d = D_u D_e^{-1} - D_u . \tag{9}$$

Substituting (9) in (4),

$$\rho = \delta^{-\frac{1}{2}} (r - D_u + D_u D_e^{-1}) \delta^{-\frac{1}{2}} , \tag{10}$$

or, remembering that $D_e = D_b D_a^{-1}$, we have from (10)

$$\rho = \delta^{-\frac{1}{2}} (r - D_u + D_u D_a D_b^{-1}) \delta^{-\frac{1}{2}} . \tag{11}$$

We let B be a vector of the beta regression weights for the tests of altered length so that

$$\rho B = \rho_c . \tag{12}$$

Substituting from (2) and (11) in (12)

$$\delta^{-\frac{1}{2}} (r - D_u + D_u D_a D_b^{-1}) \delta^{-\frac{1}{2}} B = \delta^{-\frac{1}{2}} r_c . \tag{13}$$

It can be proved that the multiple correlation for the tests of altered length is given by premultiplying (13) by B' so that

$$B' \delta^{-\frac{1}{2}} (r - D_u + D_u D_a D_b^{-1}) \delta^{-\frac{1}{2}} B = B' \delta^{-\frac{1}{2}} r_c = R_b^2 . \tag{14}$$

Now let

$$B = \delta^{\frac{1}{2}} \beta \tag{15}$$

and substitute in (14):

$$\beta'(r - D_u + D_u D_a D_b^{-1})\beta = \beta' r_c = R_b^2 . \tag{16}$$

Our problem now is to determine the D_b matrix in (16) so as to maximize R_b^2. From (16) we get

$$\beta'(r - D_u)\beta + \beta'(D_u D_a D_b^{-1})\beta = R_b^2 . \tag{17}$$

We specify that the sum of the b's shall be a constant T so that

$$T = 1' D_b 1 , \tag{18}$$

where *1* is a vector all of whose elements are unity.

Letting λ be the Lagrangian multiplier, we write

$$R_b^2 + \lambda T = \psi , \tag{19}$$

or substituting from (17) and (18) in (19)

$$\psi = \beta'(r - D_u)\beta + \beta'(D_u D_a D_b^{-1})\beta + \lambda\, 1' D_b 1 . \tag{20}$$

Formally we should now differentiate (20) partially with respect to each of the b's and equate to zero in order to obtain values for the b's which will give a maximum for (17) subject to the condition (18).* But since the first term on the right-hand side of (20) is independent of the b's, we can consider another function

$$\phi = \psi - \beta'(r - D_u)\beta$$

and write in scalar notation

$$\phi = \frac{\beta_1^2\, u_1\, a_1}{b_1} + \frac{\beta_2^2\, u_2\, a_2}{b_2} + \cdots + \lambda(b_1 + b_2 + \cdots) . \tag{21}$$

Differentiating (21) partially with respect to the b's, we have

$$\begin{aligned} \frac{\partial \psi}{\partial b_1} &= \frac{-\beta_1^2\, u_1 a_1}{b_1^2} + \lambda , \\ \frac{\partial \psi}{\partial b_2} &= \frac{-\beta_2^2\, u_2\, a_2}{b_2^2} + \lambda , \\ &\text{etc.} \end{aligned} \tag{22}$$

We have then, equating (22) to zero

*Osgood, William F. Advanced calculus. New York: McMillan, 1928, p. 180.

$$b_1 = \frac{\beta_1 (u_1 a_1)^{\frac{1}{2}}}{\lambda^{\frac{1}{2}}}$$

$$b_2 = \frac{\beta_2 (u_2 a_2)^{\frac{1}{2}}}{\lambda^{\frac{1}{2}}} \tag{23}$$

etc.

Specifying a value $\sum b$, we have from (23)

$$\lambda^{\frac{1}{2}} = \frac{\sum \beta_i (u_i a_i)^{\frac{1}{2}}}{\sum b_i}, \tag{24}$$

or in matrix notation, where *1* is a vector whose elements are all unity,

$$\lambda^{\frac{1}{2}} = \frac{1' (D_u D_a)^{\frac{1}{2}} \beta}{1' D_b 1}. \tag{25}$$

Letting D_β be a diagonal matrix of the β's, we can now write

$$D_b = D_\beta D_u^{\frac{1}{2}} D_a^{\frac{1}{2}} \frac{1' D_b 1}{1' (D_u D_a)^{\frac{1}{2}} \beta}. \tag{26}$$

From (12) and (15)

$$(r - D_u + D_u D_a D_b^{-1}) \beta = r_c . \tag{27}$$

From (26) and (27)

$$(r - D_u) \beta + D_u D_a D_u^{-\frac{1}{2}} D_a^{-\frac{1}{2}} D_\beta^{-1} \beta \frac{(1' (D_u D_a)^{\frac{1}{2}} \beta)}{1' D_b 1} = r_c. \tag{28}$$

But

$$D_\beta^{-1} \beta = 1 . \tag{29}$$

Substituting (29) in (28),

$$(r - D_u) \beta + \frac{(D_u D_a)^{\frac{1}{2}} 1 1' (D_u D_a)^{\frac{1}{2}}}{1' D_b 1} \beta = r_c \tag{30}$$

or

$$\left(r - D + \frac{(D_u D_a)^{\frac{1}{2}} 1 1' (D_u D_a^{\frac{1}{2}})}{1' D_b 1} \right) \beta = r_c. \tag{31}$$

Solving (3) for β,

$$\beta = \left(r - D_u + \frac{(D_u D_a)^{\frac{1}{2}} 1 1' (D_u D_a)^{\frac{1}{2}}}{1' D_b 1} \right)^{-1} r_c. \tag{32}$$

Equation (31) is the generalized matrix equation corresponding to equations (1) of Section I. Equation (32) simply shows the formal solution for β.

Equation (26) is the generalized matrix solution for the b's corresponding to equations (2) of Section I.

Finally, the matrix solution for R_b^2 is

$$R_b^2 = \beta' r_c \tag{33}$$

and corresponds to equation (3) of Section I.

PSYCHOMETRIKA—VOL. 14, NO. 2
JUNE, 1949

NOTE ON THE COMPUTATION OF THE INVERSE OF A TRIANGULAR MATRIX

BENJAMIN FRUCHTER
AIR FORCE TRAINING COMMAND
3309TH RESEARCH & DEVELOPMENT GROUP
UNIVERSITY OF TEXAS SUB-UNIT

A simplified method of computing the inverse of a triangular matrix is presented. It is useful with the multiple-group method of factoring the correlation matrix as well as with other factor-analysis and multiple-correlation problems.

The computation of the inverse of a triangular matrix arises in several multiple-correlation and factor-analysis problems. One important application is the multiple-group method of factoring the correlation matrix (1). In this method a table of cosines of the angular separations of the oblique axes is obtained. Thurstone, in his example of extracting three factors simultaneously, (1, 75), gives the table of cosines of the angular separations of the axes (R_{pq}) as

	p_1	p_2	p_3	
p_1	1.000	.467	.342	
p_2	.467	1.000	.437	$= R_{pq}$.
p_3	.342	.437	1.000	

It is then desired to find the matrix which will transform the oblique factors to an orthogonal frame of reference. This is accomplished by factoring matrix R_{pq} by the diagonal method. The resulting matrix is (F_{pm}):

	I	II	III	
p_1	1.000	0	0	
p_2	.467	.884	0	$= F_{pm}$.
p_3	.342	.314	.886	

The inverse of the transpose of matrix F_{pm} is the desired transformation matrix.

Several methods have been proposed for calculating the inverse of a matrix (2, 3). They are laborious and limit the convenience of the multiple-group method when a large number of factors are to be

extracted simultaneously. For a triangular matrix, however, the process can be simplified and the computation of the inverse of relatively large matrices is not prohibitive.

Computational Procedure

By way of illustration, the inverse of matrix F_{pm} above will be computed.

The inverse of the transpose of this matrix may be found directly from F_{pm} as follows:

Represent matrix F_{pm} by matrix A:

	I	II	III
p_1	1.0*	0	0
p_2	a_{21}	a_{22}	0
p_3	a_{31}	a_{32}	a_{33}

*The value in cell a_{11} is always 1.0 for this type of problem.

Represent the inverse of this matrix $(F_{pm})^{-1}$ by A^{-1}

	I	II	III
p_1	x_{11}	x_{12}	x_{13}
p_2	x_{21}	x_{22}	x_{23}
p_3	x_{31}	x_{32}	x_{33}

Then $A \cdot A^{-1} = \mathrm{I}$, or, written out:

$$\begin{array}{c|ccc} & \mathrm{I} & \mathrm{II} & \mathrm{III} \\ \hline p_1 & 1.0 & 0.0 & 0.0 \\ p_2 & a_{21} & a_{22} & 0.0 \\ p_3 & a_{31} & a_{32} & a_{33} \end{array} \cdot \begin{array}{ccc} \mathrm{I} & \mathrm{II} & \mathrm{III} \\ \hline x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{array} = \begin{array}{ccc} \mathrm{I} & \mathrm{II} & \mathrm{III} \\ \hline 1.0 & 0.0 & 0.0 \\ 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 1.0 \end{array} . \tag{1}$$

Performing the row by column matrix multiplications gives the following equations:

(Row 1 $\times$ column 1)

$$1.0\, x_{11} + 0.0\, x_{21} + 0.0\, x_{31} = 1.0 \tag{2}$$

$$x_{11} = \frac{1.0}{1.0} = 1.0 . \tag{3}$$

(Row 2 $\times$ column 1)

$$a_{21}x_{11} + a_{22}x_{21} + 0.0\, x_{31} = 0 . \tag{4}$$

As was shown in equation (3), $x_{11} = 1.0$. Hence

$$a_{21} + a_{22}x_{21} = 0 \tag{5}$$

$$x_{21} = -\frac{a_{21}}{a_{22}} = -\frac{.467}{.884} = -.529. \tag{6}$$

(Row 3 × column 1)

$$a_{31}x_{11} + a_{32}x_{21} + a_{33}x_{31} = 0\,. \tag{7}$$

From equation (3) $x_{11} = 1.0$, and from equation (6)

$$x_{21} = -\frac{a_{21}}{a_{22}}.$$

Substituting and transposing

$$a_{33}x_{31} = -a_{31} + \frac{a_{32}a_{21}}{a_{22}} \tag{8}$$

and

$$x_{31} = -\frac{a_{31}}{a_{33}} + \frac{a_{32}a_{21}}{a_{22}a_{33}} = -\frac{.342}{.886} + \frac{.314 \times .467}{.884 \times .886} = -.386 + .187 = -.199. \tag{9}$$

(Row 1 × column 2)

$$x_{12} = 0.0\,. \tag{10}$$

(Row 2 × column 2)

$$a_{21}x_{12} + a_{22}x_{22} + 0.0\,x_{32} = 1.0\,. \tag{11}$$

From equation (10) $x_{12} = 0$; hence

$$x_{22} = \frac{1}{a_{22}} = \frac{1.0}{.884} = 1.131. \tag{12}$$

(Row 3 × column 2)

$$a_{31}x_{12} + a_{32}x_{22} + a_{33}x_{32} = 0\,. \tag{13}$$

From equation (10) $x_{12} = 0$, and from equation (12)

$$x_{22} = \frac{1}{a_{22}}.$$

Substituting

$$\frac{a_{32}}{a_{22}} + a_{33}x_{32} = 0\,, \tag{14}$$

$$x_{32} = -\frac{a_{32}}{a_{22}a_{33}} = -\frac{.314}{.884 \times .886} = -.401\,. \tag{15}$$

(Row 1 × column 3)

$$x_{13} = 0.0 \tag{16}$$

(Row 2 × column 3)

$$x_{23} = 0.0 \tag{17}$$

(Row 3 × column 3)

$$a_{31}x_{13} + a_{32}x_{23} + a_{33}x_{33} = 1.0\,. \tag{18}$$

From equations (16) and (17), $x_{13} = 0$ and $x_{23} = 0$.

Therefore

$$x_{33} = \frac{1}{a_{33}} = \frac{1}{.886} = 1.129\,. \tag{19}$$

Putting the results in the form of a matrix gives

	I	II	III	
p_1	1.000	.000	.000	
p_2	—.529	1.131	.000	$= (F_{pm})^{-1}$.
p_3	—.199	—.401	1.129	

This is the inverse of F_{pm}. Comparing matrices F_{pm} and $(F_{pm})^{-1}$, it may be observed that wherever a zero occurs above the principal diagonal in the former it also occurs in the latter. The values along the principal diagonal of $(F_{pm})^{-1}$ are the reciprocals of the corresponding values of F_{pm}. The other values of $(F_{pm})^{-1}$ are obtained by means of simple equations similar to those outlined above.

The desired inverse is merely its transpose $(F'_{pm})^{-1}$ and is written

	I	II	III	
p_1	1.000	—.529	—.199	
p_2	.000	1.131	—.401	$= (F'_{pm})^{-1}$.
p_3	.000	.000	1.129	

REFERENCES

1. Thurstone, L. L. A multiple group method of factoring the correlation matrix. *Psychometrika*, 1945, **10**, 73-78.
2. Thurstone, L. L. Multiple-factor analysis. Chicago: University of Chicago Press, 1947, 46-48.
3. Deemer, Walter L. (Ed.) *Records, analysis, and test procedures.* Army Air Force Aviation Psychology Program Research Reports. Washington: U. S. Government Printing Office, 1947, 501-506.

PSYCHOMETRIKA—VOL. 14, NO. 1
MARCH, 1949

A METHOD OF MATRIX ANALYSIS OF GROUP STRUCTURE

R. DUNCAN LUCE AND ALBERT D. PERRY
GRADUATE STUDENTS, DEPARTMENT OF MATHEMATICS
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Matrix methods may be applied to the analysis of experimental data concerning group structure when these data indicate relationships which can be depicted by line diagrams such as sociograms. One may introduce two concepts, n-chain and clique, which have simple relationships to the powers of certain matrices. Using them it is possible to determine the group structure by methods which are both faster and more certain than less systematic methods. This paper describes such a matrix method and applies it to the analysis of practical examples. At several points some unsolved problems in this field are indicated.

1. *Introduction*

In a number of branches of the social sciences one encounters problems of the analysis of relationships between the elements of a group. Frequently the results of these investigations may be presented in diagrammatic form as sociograms, organization charts, flow charts, and the like. When the data to be analyzed are such that a diagram of this type may be drawn, the analysis and presentation of the results may be greatly expedited by using matrix algebra. This paper presents some of the results of an investigation of this application of matrices. Initial trials in the determination of group structures indicate that the matrix method is not only faster but also less prone to error than manual investigation.*

The second section of this paper presents certain concepts used in the analysis and associates matrices with the group in question. The third states the results obtained and the fourth gives illustrations of their application. Finally, section five contains a mathematical formulation of the theory and derivation of the results presented in section three.

2. *Definitions*

2.01. The types of relationships which this method will handle are: man a chooses man b as a friend, man a commands man b, a sends messages to b, and so forth. Since in a given problem we concern

*Some of these examples have been worked out by the Research Center for Group Dynamics, Massachusetts Institute of Technology, in conjunction with some of its research.

ourselves with one sort of relation, no confusion arises from replacing the description of the relationship by a symbol $=>$. Thus, instead of "man i chooses man j as a friend," we write "$i => j$." If, on the other hand, man i had not chosen man j, we would have written "$i \neq > j$," using the symbol $\neq >$ to indicate the negation of the relationship denoted by $=>$.

2.02. Situations such as mutual choice of friends or two-way communication would thus be indicated by $i => j$ and $j => i$, or briefly, $i <=> j$. We describe such situations by saying that a *symmetry* exists between i and j.

2.03. When the choice is not mutual, that is $i => j$ or $j => i$ but not both, we say an *antimetry* exists between i and j.

2.04.* The data to be analyzed are presented in a matrix X as follows: the i,j entry (x_{ij}) has the value of 1 if $i => j$ and the value 0 if $i \neq > j$. For convenience we place the main diagonal terms equal to zero, i.e., $x_{ii} = 0$ for all i. This convention, $i \neq > i$, does not restrict the applicability of the method, since there is little significance in such statements as "Jones chooses himself as a friend."

Suppose, for example, that we had a group of four members with the following relationships: $a => b$, $b => a$, $b => d$, $d => b$, $c => a$, $c => b$, $d => a$, and $d => c$. All other possible combinations of $a,b,c,$ and d are related by the symbol $\neq >$. The X matrix associated with this group is:

$$\begin{array}{c} \\ a \\ b \\ c \\ d \end{array}\begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

2.05. From the X matrix we extract a symmetric matrix S having entries s_{ij} determined by $s_{ij} = s_{ji} = 1$ if $x_{ij} = x_{ji} = 1$, and otherwise $s_{ij} = s_{ji} = 0$. All the symmetries in the group are indicated in the matrix S. The S matrix associated with the above X matrix is:

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

*In the course of the present work it was brought to our attention that in "A matrix approach to the analysis of sociometric data," *Sociometry*, 1946, **9**, 340-347, Elaine Forsyth and Leo Katz have used matrices to represent sociometric relations. They considered a three-valued logic rather than the present two-valued one, and the operations on the matrices are different from the ones discussed in this paper.

To indicate the i,j entry of the matrix X^n, which is the n^{th} power of X, we shall employ the symbol $x_{ij}^{(n)}$. Similarly, the i,j entry of S^n is $s_{ij}^{(n)}$.

2.06. In the group considered above, we had $a => b$, $b => d$, and $d => c$ as three of the relations. If the symbol $=>$ indicates the relationship "sends messages to," it appears that a can send a message to c in three steps, via b and d. We call this three-step path a 3-chain from a to c. Rather than writing out the above sequence of relations, we may omit the symbol $=>$ and simply write the 3-chain as a,b,d,c.

In a group involving more elements one might have the 5-chains: a,e,c,b,d,f and a,d,b,c,d,e. We notice that the first sequence involves five steps between six elements of the group. The second sequence also involves five steps but only five elements of the group, since the element d appears as both the second and fifth member of the sequence. Thus, although these two five-step sequences contain different numbers of elements of the group, they both have six members. Using this concept of membership in a sequence, an n-step sequence has $n+1$ members.

These examples of 3-chains and 5-chains suggest a general definition for a property within the group: an ordered sequence with $n+1$ members, $i, a, b, \cdots, p, q, j$, is an n-*chain* from i to j if and only if

$$i => a, a => b, \cdots, p => q, q => j.$$

2.07. When two n-chains have the same elements in the same order, i.e., the same members, then they are said to be *equal,* and otherwise they are *distinct.* It is important in this definition of equality that it be recognized that both the elements of the group and their order in the sequence are considered. The two chains i,j,k,l,p and i,p,k,j,l are distinct though they contain the same five elements.

2.08. When the same element occurs more than once in an n-chain, the n-chain is said to be *redundant.* (Thus, in a group of m elements any n-chain with n greater than m is redundant). The chains a,b,e,d,b,c and a,c,a,b,d,c,e are, for example, both redundant, for the element b occurs twice in the former and the elements a and c both occur twice in the latter. An example of a non-redundant 5-chain is a,d,p,b,q,j.

2.09. A subset of the group forms a *clique* provided that it consists of three or more members each in the symmetric relation to each other member of the subset, and provided further that there can be found no element outside the subset that is in the symmetric relation to each of the elements of the subset. The application of this definition to the concept of friendship is immediate: it states that a

* equal to

set of more than two people form a clique if they are all mutual friends of one another. In addition, the definition specifies that subsets of cliques are not cliques, so that in a clique of five friends we shall not say that any three form a clique. Although the word "clique" immediately suggests friendship, the definition is useful in the study of other relationships.

2.10. This definition of clique has two possible weaknesses: first, if each element of the group is related by $=>$ to no more than c other elements of the group, then we can detect only cliques with at most $c+1$ members; and second, there may exist within the group certain tightly knit subgroups which by the omission of a few symmetries fail to satisfy the definition of a clique but which nonetheless would be termed, non-technically, "cliques." It may be possible to alleviate these difficulties by the introduction of so called "n-cliques" which comprise the set of n elements which form two distinct n-chains from each element of the set to itself. This requires that the n-chains be redundant with the only recurring element being the end-point and also that all the relations in the n-chains be symmetric.

This definition means that the four elements $a,b,c,$ and d form a 4-clique if the 4-chains (for example) a,b,c,d,a and $a,d,c,b,a,$ both exist. These by the definition of n-chain require that the relations

$$a <=> b\,,\ b <=> c\,,\ c <=> d\,,\ d <=> a$$

exist, but nothing is said about the relations between a and c, and b and d. The original definition requires, in addition, that

$$a <=> c \quad \text{and} \quad b <=> d$$

for $a,b,c,$ and d to form a clique of four members. Thus we see that the definition of n-clique considers "circles" of symmetries, but it fails to consider the symmetric "cross" terms that exist between the members of the n-clique. These cross terms will be investigated, however, by determining whether any m of these n-elements form an m-clique.

The usefulness of the definition of n-clique can be judged only after experience has been gained in its application. This is not conveniently possible at present, unfortunately, because the problem of the general determination of redundant n-chains has not been solved (see §5.09).

The most general definition of a clique-like structure including antimetries will not be discussed, for it is believed that this will not be amenable to a concise mathematical formulation.

3. *Statement of Results*

3.01. In X^n the entry $x_{ij}^{(n)} = c$ if and only if there are c dis-

tinct n-chains from i to j (for proof see §5.04). Thus, if in the fifth power of a matrix of data X we find that the number 9 occurs in the third row of the seventh column, we may conclude that there are 9 distinct 5-chains from element 3 to element 7.

3.02. In X^2 the i^{th} main diagonal entry has the value m if and only if i is in the symmetric relation with m elements of the group (§5.05). Since by the definition of a clique each element i in a clique of t members must be in the symmetric relation to each of the t—1 other elements, it is necessary that $x_{ii}^{(2)} \geqslant t-1$ for i to be in a clique of t members. We may not, however, conclude from the fact that $x_{ii}^{(2)} \geqslant t-1$ that i is necessarily contained in a clique of t members.

3.03. An element i is contained in a clique if and only if the i^{th} entry of the main diagonal of S^3 is positive (§5.06). The main diagonal terms of S^3 will be either 0 or even positive numbers in all cases, and when the value of the entry is 0 the associated element is not in a clique.

3.04. If, in S^3, t entries of the main diagonal have the value $(t-2)(t-1)$ and all other entries of the main diagonal are zero, then these t elements form a clique of t members (§5.08). It also follows from the next statement (§3.05) that if there is only one clique of t members then these t elements will have a main diagonal value in S^3 of $(t-2)(t-1)$. The former statement is, however, the more significant in analysis, for it is the aim to go from the matrix representation to the group structure. There is no difficulty in going from the structure to the matrices.

3.05. Since by statement 3.03 the main diagonal values of S^3 are dependent only on the clique structure of the group, it is to be expected that a formula relating these values and the clique structure is possible. If an element i is contained in m different cliques each having t_ν members, and if there are d_k elements common to the k^{th} clique and all the preceding ones, then

$$s_{ii}^{(3)} = \sum_{\nu=1}^{m} \{(t_\nu - 2)(t_\nu - 1) - (d_\nu - 2)(d_\nu - 1)\} + 2$$

(§5.07). Thus, if we have three cliques: (5,7,9,10), (1,4,9), and (1,2,5,9,11), then $d_1 = 0$, for there are no preceding cliques; $d_2 = 1$, for only element 9 is common to the second and first cliques; and $d_3 = 3$, for clique three has the elements 1,5, and 9 common with the first two cliques. Substituting $t_1 = 4$, $t_2 = 3$, $t_3 = 5$, $d_1 = 0$, $d_2 = 1$, and $d_3 = 3$ and evaluating the formula for element 9, which is the only one common to all three cliques, we obtain

$$\begin{aligned} s_{99}{}^{(3)} &= [(4-2)(4-1)-(0-2)(0-1)] \\ &\quad + [(3-2)(3-1)-(1-2)(1-1)] \\ &\quad + [(5-2)(5-1)-(3-2)(3-1)]+2 \\ &= 18\,. \end{aligned}$$

In the evaluation of this formula it is immaterial how the cliques are numbered initially; however, it is essential once the numbering is chosen that we be consistent.

3.06. The redundant 2-chains of a matrix X are the main diagonal entries of X^2 (§5.09). Thus for a matrix

$$X = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

with the square

$$X^2 = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 2 & 0 & 0 & 0 & 2 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 & 0 \end{bmatrix},$$

the matrix of redundant 2-chains is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

To obtain the matrix of redundant 3-chains we compute the following matrix, in which the symbol $R^{(2)}$ stands for the matrix of redundant 2-chains:

$$XR^{(2)} + R^{(2)}X - S\,.$$

Deleting in this sum the main diagonal and replacing it by the main diagonal of X^3 gives the matrix of redundant 3-chains (§5.09). If the main diagonal of $XR^{(2)} + R^{(2)}X - S$ is denoted by $Y^{(3)}$ and the main diagonal of $X^{(3)}$ by $Z^{(3)}$, then let $E^{(3)} = Z^{(3)} - Y^{(3)}$ and thus the matrix of redundant 3-chains, $R^{(3)}$, is given by

$$R^{(3)} = XR^{(2)} + R^{(2)}X + E^{(3)} - S\,.$$

It has not yet been possible to develop formulas which will give the matrix of redundant n-chains for n larger than 3. What work that has been done in this direction is presented in §5.09.

3.07. The several theorems on cliques give a method that to some extent determines the clique structure independent of the rest of the group structure. It would be desirable to find a simple scheme that determines the clique structure directly. Since a certain amount of knowledge in this direction can be obtained from S^3, it is conjectured that possibly there is a simple formula relating clique structure to the numbers in S^3. As yet no such formula has been developed.

In a consideration of this problem, it was questioned whether certain aspects of the structure would be lost in the multiplication, which, if true, might make the discovery of the desired formula impossible. The following theorem shows that neither the clique structure nor any of the properties of S are lost in the matrix S^3: Any real symmetric matrix has one and only one real symmetric n^{th} root if n is a positive odd integer (§5.12). This theorem is somewhat more general than was required, since it does not restrict the entries in the n^{th} root to 0 and 1, and since it is true for any odd root rather than just the cube root. (In general the real symmetric even roots are not unique.)

This theorem suggests a further problem to be solved: to find a symmetric group structure which will insure the presence of certain prescribed minimum n-chain conditions for odd n. To carry this out it will probably prove necessary to discover a theorem that uses not only the realness and symmetry of the S matrix and its powers, but in addition the fact that only the numbers 1 and 0 may be entries in S.

4. *Examples*

4.01. As the first example, let us compare and analyze the friendship structure in the two following hypothetical groups. The matrices are (where a blank entry indicates a zero):

$$
\begin{array}{cc}
 & \text{I} \\
 & \begin{matrix}1&2&3&4&5&6&7&8&9&10\end{matrix} \\
\begin{matrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\end{matrix} &
\begin{bmatrix}
 & 1 & & 1 & & & 1 & 1 & & 1 \\
1 & & & 1 & 1 & & 1 & 1 & & 1 \\
1 & & & & & & & 1 & & \\
1 & 1 & & & & & 1 & & & 1 \\
 & & & & & 1 & & & & \\
 & & & & 1 & & & & & \\
1 & 1 & & 1 & & & & 1 & & \\
1 & 1 & & 1 & & & 1 & & 1 & \\
 & & & & & & & & & \\
1 & 1 & & 1 & 1 & & & & &
\end{bmatrix}
\end{array}
\qquad
\begin{array}{cc}
 & \text{II} \\
 & \begin{matrix}1&2&3&4&5&6&7&8&9&10\end{matrix} \\
\begin{matrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\end{matrix} &
\begin{bmatrix}
 & & 1 & 1 & 1 & & & & & 1 \\
 & & & & & 1 & & & 1 & \\
1 & & & 1 & & & & & & \\
1 & & 1 & & & 1 & & & 1 & 1 \\
1 & & & & & & & & & \\
 & 1 & & & & & 1 & & & \\
 & 1 & & & & 1 & & & & \\
 & & & & 1 & & & & 1 & \\
 & & & 1 & 1 & & & 1 & & \\
1 & & 1 & & & & 1 & & &
\end{bmatrix}
\end{array}
$$

The associated S matrices are:

$$
\begin{array}{cc}
 & \begin{matrix}1&2&3&4&5&6&7&8&9&10\end{matrix} \\
\begin{matrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\end{matrix} &
\begin{bmatrix}
 & 1 & & 1 & & & 1 & 1 & & 1 \\
1 & & & 1 & & & 1 & 1 & & 1 \\
 & & & & & & & & & \\
1 & 1 & & & & & 1 & & & 1 \\
 & & & & & 1 & & & & \\
 & & & & 1 & & & & & \\
1 & 1 & & 1 & & & & 1 & & \\
1 & 1 & & & & & 1 & & & \\
 & & & & & & & & & \\
1 & 1 & & 1 & & & & & &
\end{bmatrix}
\end{array}
\qquad
\begin{array}{cc}
 & \begin{matrix}1&2&3&4&5&6&7&8&9&10\end{matrix} \\
\begin{matrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\end{matrix} &
\begin{bmatrix}
 & & 1 & 1 & 1 & & & & & 1 \\
 & & & & & 1 & & & & \\
1 & & & 1 & & & & & & \\
1 & & 1 & & & & & & 1 & \\
1 & & & & & & & & & \\
 & 1 & & & & & 1 & & & \\
 & & & & & 1 & & & & \\
 & & & & & & & & 1 & \\
 & & & 1 & & & & 1 & & \\
1 & & & & & & & & &
\end{bmatrix}
\end{array}
$$

The S^2 matrices are:

$$
\begin{array}{cc}
 & \text{I} \\
 & \begin{matrix}1&2&3&4&5&6&7&8&9&10\end{matrix} \\
\begin{matrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\end{matrix} &
\begin{bmatrix}
5 & 4 & & 3 & & & 3 & 2 & & 2 \\
4 & 5 & & 3 & & & 3 & 2 & & 2 \\
 & & & & & & & & & \\
3 & 3 & & 4 & & & 2 & 3 & & 2 \\
 & & & & 1 & & & & & \\
 & & & & & 1 & & & & \\
3 & 3 & & 2 & & & 4 & 2 & & 3 \\
2 & 2 & & 3 & & & 2 & 3 & & 2 \\
 & & & & & & & & & \\
2 & 2 & & 2 & & & 3 & 2 & & 3
\end{bmatrix}
\end{array}
\qquad
\begin{array}{cc}
 & \text{II} \\
 & \begin{matrix}1&2&3&4&5&6&7&8&9&10\end{matrix} \\
\begin{matrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\end{matrix} &
\begin{bmatrix}
4 & & 1 & 1 & & & & & 1 & \\
 & 1 & & & & & 1 & & & \\
1 & & 2 & 1 & 1 & & & & 1 & 1 \\
1 & & 1 & 3 & 1 & & & 1 & & 1 \\
 & & 1 & 1 & 1 & & & & & 1 \\
 & & & & & 2 & & & & \\
 & 1 & & & & & 1 & & & \\
 & & & 1 & & & & 1 & & \\
1 & & 1 & & & & & & 2 & \\
 & & 1 & 1 & 1 & & & & & 1
\end{bmatrix}
\end{array}
$$

Here the differences between the groups are becoming evident. In group I, men 3 and 9 have no mutual friends, since $s_{33}^{(2)} = s_{99}^{(2)} = 0$

(§3.02). Thus, as far as symmetric relationships are concerned, these men are isolated from the group. In the same way we determine that 5 and 6 each have just one symmetric friendship relation ($s_{55}^{(2)} = s_{66}^{(2)} = 1$, §3.02) which we determine to be 5 <=> 6 from the S matrix. The remaining elements in S^2 form a rather dense set of quite large numbers, which means, roughly, a tightly knit group.

In the second group, on the other hand, every man has a non-zero main diagonal in S^2. The men 2, 5, 7, 8, and 10 each have a single mutual friend, which we determine to be: 2 <=> 6, 5 <=> 1, 7 <=> 6, 8 <=> 9, and 10 <=> 1. Then since $s_{66}^{(2)} = 2$ and since we have just cited 6's two mutual friends, 6 need not be considered further. We note that the off-diagonal areas of this S^2 matrix are not so completely filled as group I, indicating that the group is not so tightly bound.

The S^3 matrices indicate the differences in compactness of the structures quite clearly:

$$\text{I} \qquad \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \left[\begin{array}{cccccccccc} 14 & 15 & & 14 & & & 14 & 12 & & 12 \\ 15 & 14 & & 14 & & & 14 & 12 & & 12 \\ & & & & & & & & & \\ 14 & 14 & & 10 & & & 13 & 8 & & 10 \\ & & & & & 1 & & & & \\ & & & & 1 & & & & & \\ 14 & 14 & & 13 & & & 10 & 10 & & 8 \\ 12 & 12 & & 8 & & & 10 & 6 & & 7 \\ & & & & & & & & & \\ 12 & 12 & & 10 & & & 8 & 7 & & 6 \end{array}\right] \end{array}$$

$$\text{II} \qquad \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\ \left[\begin{array}{cccccccccc} 2 & & 5 & 6 & 4 & & & 1 & 1 & 4 \\ & & & & & 2 & & & & \\ 5 & & 2 & 4 & 1 & & & 1 & 1 & 1 \\ 6 & & 4 & 2 & 1 & & & & 4 & 1 \\ 4 & & 1 & 1 & & & & & 1 & \\ & 2 & & & & & 2 & & & \\ & & & & & 2 & & & & \\ 1 & & 1 & & & & & & 2 & \\ 1 & & 1 & 4 & 1 & & & 2 & & 1 \\ 4 & & 1 & 1 & & & & & 1 & \end{array}\right] \end{array}$$

Since the corresponding main diagonal terms are non-zero, men 1, 2, 4, 7, 8, and 10 of group I are in cliques (§3.03). These, with 3 and 9 which have no symmetries in the group and 5 and 6 which are mutual friends, account for all members of the group. The terms $s_{88}^{(3)} = s_{1010}^{(3)} = 6$ suggest a clique of four members; however, the existence of other main diagonal terms makes it impossible to apply the formula $(t-2)(t-1)$ (§3.04). Investigating in S first the elements 1, 2, and 4 because their columns have the largest values in the tenth row, we find that elements 1, 2, 4, and 10 form a clique of four members. In the eighth row the largest entries are in columns 1, 2, and 7, and an investigation reveals that 1, 2, 7, and 8 form a clique of four men, which then overlaps the first clique by the men 1 and 2. In row four the largest entries are found in columns 1, 2, and 7. We then find that 1, 2, 4, and 7 form a clique of four elements which

overlaps the previous two. All the men contained in cliques have been accounted for at least once, and a check either with the formula for main diagonal entries (§3.05) or directly in the S matrix indicates that all the cliques have been discovered. This, coupled with what we discovered in S^2, completely determines the symmetric structure of the first group.

For purposes of qualitative judgment and a guide to carrying out analysis, we note that the first two rows of S^3 present an interesting summary of the clique structure. The entries $s_{12}^{(3)}$ and $s_{21}^{(3)}$ have the largest values, next largest are in columns four and seven, and then finally in columns eight and ten. Men 1 and 2 are contained in all three cliques, 4 and 7 are each contained in two cliques, and finally men 8 and 10 are each in only one clique. This indicates that the magnitude of the off-diagonal terms determines to some extent the amount and structural position of the overlap of cliques.

In group II there are only three elements with non-zero main-diagonal entries, all with the value 2. This fits the formula $(t—2)(t—1)$ with $t = 3$ (§3.04). Thus the men 1, 3, and 4 form a clique of three members. Returning to S^2, we see that there remains one unaccounted symmetry each for men 4 and 9, hence $4 <=> 9$.

In group I, the off-diagonal terms are large in magnitude and are quite dense in the array, with some rows completely empty or with single entries in the S^3 matrix. This indicates a closely knit group with certain men definitely excluded. The S^3 matrix for the second group has fewer entries of a smaller value indicating a less tightly knit structure, but it has no empty rows and only one row with a single entry; that is, it has fewer people than group I who are not accepted by the group or who do not accept it.

A consideration of the matrix $X — S$ will give all the antimetries in the groups and complete the analysis of the structures.

It is clear that this procedure gains strength as the complexity of the problem increases, for the analysis of a twenty-element group is little more difficult than that of a ten-element group.

4.02. The second example is a communication system comprising two-way links between seven stations such as might occur in a telephone or telegraph circuit. The number of channels of a given number of steps (i.e., n-chains in the general theory) between any two points and the minimum number of steps required to complete contact between two stations will be determined. Suppose the matrix of one-step contacts is:

$$
\begin{array}{c}
\\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix}
 & 1 & 1 & & & & 1 \\
1 & & 1 & & & & \\
1 & 1 & & 1 & 1 & & \\
 & & 1 & & 1 & 1 & 1 \\
 & & 1 & 1 & & 1 & 1 \\
 & & & 1 & 1 & & 1 \\
1 & & & 1 & 1 & 1 &
\end{bmatrix}
\end{array}
$$

which in this case is also the S matrix. Then two-step connections are given by X^2:

$$
\begin{array}{c}
\\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix}
3 & 1 & 1 & 2 & 2 & 1 & 0 \\
1 & 2 & 1 & 1 & 1 & 0 & 1 \\
1 & 1 & 4 & 1 & 1 & 2 & 3 \\
2 & 1 & 1 & 4 & 3 & 2 & 2 \\
2 & 1 & 1 & 3 & 4 & 2 & 2 \\
1 & 0 & 2 & 2 & 2 & 3 & 2 \\
0 & 1 & 3 & 2 & 2 & 2 & 4
\end{bmatrix}
\end{array}
$$

and the three-step ones by X^3:

$$
\begin{array}{c}
\\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{bmatrix}
2 & 4 & 8 & 4 & 4 & 4 & 8 \\
4 & 2 & 5 & 3 & 3 & 3 & 3 \\
8 & 5 & 4 & 10 & 10 & 5 & 5 \\
4 & 3 & 10 & 8 & 9 & 9 & 11 \\
4 & 3 & 10 & 9 & 8 & 9 & 11 \\
4 & 3 & 5 & 9 & 9 & 6 & 8 \\
8 & 3 & 5 & 11 & 11 & 8 & 6
\end{bmatrix}
\end{array}
$$

From the former, the two connections $1 \overset{(2)}{<=>} 7$ and $2 \overset{(2)}{<=>} 6$ cannot be realized because $x_{17}^{(2)} = x_{71}^{(2)} = 0$ and $x_{26}^{(2)} = x_{62}^{(2)} = 0$ (§3.01). The contacts are possible in three steps, however, since X^3 is completely filled. Thus two steps are sufficient for most contacts and three steps for all.

In determining the number of paths between two points it is desirable to eliminate redundant paths. For two-step communication this is done by deleting the main diagonal of X^2. The remaining terms represent the number of two-step paths between the stations indicated. The matrix of redundancies for three-step communication is given by $R^{(3)} = XR^{(2)} + R^{(2)}X + E^{(3)} - S$ (§3.06), which works out to be:

$$\begin{array}{cc}
 & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} &
\begin{bmatrix}
2 & 4 & 6 & & & & 6 \\
4 & 2 & 5 & & & & \\
6 & 5 & 4 & 7 & 7 & & \\
 & & 7 & 8 & 7 & 6 & 7 \\
 & & 7 & 7 & 8 & 6 & 7 \\
 & & & 6 & 6 & 5 & 6 \\
6 & & & 7 & 7 & 6 & 6
\end{bmatrix}
\end{array}$$

The matrix of non-redundant three-step communication paths is $X^3 - R^{(3)}$:

$$\begin{array}{cc}
 & \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} &
\begin{bmatrix}
 & & 2 & 4 & 4 & 4 & 2 \\
 & & & 3 & 3 & 3 & 3 \\
2 & & & 3 & 3 & 5 & 5 \\
4 & 3 & 3 & & 2 & 3 & 4 \\
4 & 3 & 3 & 2 & & 3 & 4 \\
4 & 3 & 5 & 3 & 3 & & 2 \\
2 & 3 & 5 & 4 & 4 & 2 &
\end{bmatrix}
\end{array}$$

We notice that the three-step paths between 1 and 2 and 2 and 3 are all redundant but that there are two-step paths for these combinations. All other combinations have at least two three-step paths joining them.

5. *Mathematical Theory*

5.01 To carry out the following mathematical formulation and the proofs of theorems it is convenient to use some of the symbolism and nomenclature of point set theory. As there is some diversity in the literature, the symbols used are:

Sets are either defined by enumeration or by properties of the elements of the set in the form: symbol for the set [symbols used for elements of the set | defining properties of these elements]. When a single element i is treated as a set it will be denoted by (i), otherwise sets will be denoted by upper case Greek letters.

The intersection of (elements common to) two sets Γ and Φ is denoted by $\Gamma \cdot \Phi$.

The union of two sets Γ and Φ (elements contained in either or both) is denoted by $\Gamma + \Phi$. The context will make it clear whether the symbol + refers to addition, matrix addition, or union.

The inclusion of a set Γ in another set Φ (all elements of Γ are elements of Φ) is denoted by $\Gamma < \Phi$. The negation is $\Gamma <^* \Phi$.

If $\Phi < \Xi$, then the complement of Φ with respect to Ξ, Φ', is de-

fined by $\Phi + \Phi' = \Xi$ and $\Phi \cdot \Phi' = 0$ where 0 is the null set.

The inclusion of a single element i in a set Φ is denoted by $i \varepsilon \Phi$.

For any two elements i and j of a set Ξ and a subset Ω of Ξ:

$(i) + (j) < \Omega$ if and only if $i \varepsilon \Omega$ and $j \varepsilon \Omega$.
$(i) + (j) <^{*} \Omega$ implies $i \varepsilon \Omega'$ and/or $j \varepsilon \Omega'$.

The symbol $\delta_{ij} \begin{array}{ll} =1 & \text{if} \quad i=j \\ =0 & \text{if} \quad i \neq j \end{array}$

5.02. Consider a finite set Ξ of x elements denoted by $1, 2, \cdots, i, \cdots, j, \cdots, x$ for which there is defined a relationship $=>$ between elements and its negation $\neq>$ having the properties:

1. Either $i => j$ or $i \neq> j$ for all i and $j \varepsilon \Xi$.
2. $i \neq> i$.

Let a number x_{ij} be associated with i and j such that

$$x_{ij} \begin{array}{ll} =1 & \text{if} \quad i => j \\ =0 & \text{if} \quad i \neq> j. \end{array}$$

A matrix $X = [x_{ij}]$ is formed from the numbers x_{ij}. It will be found useful to denote the i,j entry of the n^{th} power of X, X^n, by $x_{ij}^{(n)}$.

A *symmetry* is said to exist between i and j if and only if $i => j$ and $j => i$, in which case we may write $i <=> j$. For the matrix X this requires that $x_{ij} = x_{ji} = 1$. If, however, either $i => j$ and $j \neq> i$ or $i \neq> j$ and $j => i$ then an *antimetry* is said to exist between i and j.

The *symmetric matrix* S associated with the matrix X is defined by $S = [s_{ij}]$, where

$$s_{ij} = s_{ji} \begin{array}{ll} =1 & \text{if} \quad x_{ij} = x_{ji} = 1, \quad \text{i.e.,} \quad i <=> j. \\ =0 & \text{otherwise.} \end{array}$$

The i,j entry of the n^{th} power of S is $s_{ij}^{(n)}$.

5.03. Definitions:

1. An ordered sequence with $n+1$ members, $i \equiv \gamma_1, \gamma_2, \cdots, \gamma_n, \gamma_{n+1} \equiv j$, is an n-*chain* Γ from i to j if and only if

$$i \equiv \gamma_1 => \gamma_2, \gamma_2 => \gamma_3, \cdots, \gamma_n => \gamma_{n+1} \equiv j.$$

In brief, $i \overset{(n)}{=>} j$ indicates that there exists an n-chain from i to j, which may also be enumerated as $i \equiv \gamma_1, \gamma_2, \cdots, \gamma_n, \gamma_{n+1} \equiv j$, or, when no ambiguity will arise, as $i, k, l, \cdots, p, q, j$ with the ordering being indicated by the written order of the sequence.

2. Two n-*chains* Γ and Φ *are equal* if and only if the r^{th} member of Γ equals the r^{th} member of Φ, i.e., $\gamma_r = \phi_r$, for $1 \leqslant r \leqslant n+1$.

If this is not true, then Γ and Φ are *distinct*.

3. Each pair of elements γ_k and γ_m of an n-chain with $1 \leqslant k < m \leqslant n+1$ and $\gamma_k = \gamma_m$ is said to be the *redundant pair* (k,m). An n-chain is *redundant* if and only if it contains at least one redundant pair.

4. The elements $1, 2, \cdots, t$ $(t \geqslant 3)$ form a *clique* Θ of t members if and only if each element of Θ is symmetric with each other element of Θ, and there is no element not in Θ symmetric with all elements of Θ.

This is equivalent to

$x_{ij} = 1 - \delta_{ij}$ for $i, j = 1, 2, \cdots, t$ but not for $i, j = 1, 2, \cdots, t, t+1$, whatever the $(t+1)^{st}$ element.

5.04. Theorem 1: $x_{ij}^{(n)} = c$ if and only if there exist c distinct n-chains from i to j.
Proof: By definition of matrix multiplication

$$x_{ij}^{(n)} = \sum_{k \varepsilon \Xi} \cdots \sum_{q \varepsilon \Xi} x_{ik} x_{kl} \cdots x_{pq} x_{qj},$$

with the summations over n—1 indices. Suppose that the indices have been selected such that $i, k, l, \cdots, p, q, j$ is an n-chain from i to j.

Then by definition 1 (§5.03)

$$x_{ik} = x_{kl} = \cdots = x_{pq} = x_{qj} = 1,$$

and if the indices were not so selected then at least one $x_{rs} = 0$. Thus n-chains contribute 1 to the sum and other ordered sequences contribute 0. Since the indices take on each possible combination of values just once, every distinct n-chain is represented just once. If there are c such n-chains, then there are a total of c ones in the summation.

5.05. Theorem 2: An element of Ξ has a main diagonal value of c in X^2 if and only if it is symmetric with c elements of Ξ.
Proof: Let Φ be the set of j's for which $i <=> j$. By definition

$$x_{ii}^{(2)} = \sum_{j \varepsilon \Phi} x_{ij} x_{ji} + \sum_{j \varepsilon \Phi'} x_{ij} x_{ji} = \sum\nolimits_1 + \sum\nolimits_2 .$$

$\sum_1 = c$ by theorem 1 (§5.04) and $\sum_2 = 0$ because i and j are not symmetric for $j \varepsilon \Phi'$, so either $x_{ij} = 0$ or $x_{ji} = 0$ or both. Thus if i is symmetric with c elements of Ξ, $x_{ij}^{(2)} = c$.

If $x_{ii}^{(2)} = c$, then by theorem 1 there exist c distinct j's such that $x_{ij} = x_{ji} = 1$, i.e., $i <=> j$ for c j's.

5.06. Theorem 3: An element i is contained in a clique if and only if the i^{th} entry of the main diagonal of S^3 is positive.

Proof: Suppose that i is contained in a clique Θ.

By definition

$$s_{ii}^{(3)} = \sum_{(j)+(k)<\Xi} \sum s_{ij} s_{jk} s_{ki}.$$

Select j and k such that $(j) + (k) < \Theta$ and such that $i \neq j \neq k \neq i$. Such elements exist by the definition of a clique (definition 4, §5.03). It is true by the definition of a clique and of the matrix S that: $s_{ij} = s_{ji} = s_{jk} = s_{kj} = s_{ik} = s_{ki} = 1$ for such j and k. Thus this choice of j and k contributes 2 to the summation, and because $s_{ij} \geqslant 0$ for all i and j there are no negative contributions to the sum; therefore $s_{ii}^{(3)} \geqslant 2 > 0$.

Suppose that $s_{ii}^{(3)} > 0$. Then there exists at least one pair of elements of j and k such that $s_{ij} = s_{jk} = s_{ki} = 1$ and this implies $i <=> j$, $j <=> k$, and $k <=> i$. If there are no other elements symmetric with i, j, and k then these three form a clique. If there is another element symmetric with these three, then consider the set of four formed by adding it to the previous three. If there is no other element symmetric with these four, they form a clique. If there is, add it to the set and continue the process. Since the set Ξ contains only a finite number of elements, the process must terminate giving a clique containing i.

5.07. Theorem 4: If 1) Θ_σ are cliques of t_σ members, 2) the sets $\Delta_\nu = \Theta_\nu \cdot (\Theta_1 + \Theta_2 + \cdots + \Theta_{\nu-1})$ have d_ν members, and 3) i is contained in the cliques Θ_σ, $\sigma = 1, 2, \cdots, m$, then

$$s_{ii}^{(3)} = \sum_{\sigma=1}^{m} \{(t_\sigma - 2)(t_\sigma - 1) - (d_\sigma - 2)(d_\sigma - 1)\} + 2.$$

Proof: By definition

$$s_{ii}^{(3)} = \sum_{(j)+(k)<\Xi} \sum s_{ij} s_{jk} s_{ki}.$$

The set of all the pairs j, k is the union of the following three mutually exclusive sets:

Ψ_1 $[j, k \mid$ there exists ν such that $(j) + (k) < \Theta_\nu$; there does not exist α such that $(j) + (k) < \Theta_\alpha$, $(j) + (k) <^{\bullet} \Delta_\alpha]$

Ψ_2 $[j, k \mid$ there does not exist α such that $(j) + (k) < \Theta_\alpha]$

Ψ_3 $[j, k \mid$ there exists α such that $(j) + (k) < \Theta_\alpha$, $(j) + (k) <^* \Delta_\alpha]$.

1. For Ψ_1 then either

a) $(j) + (k) <^* \Theta_\alpha$ for all α. This is not possible because $(j) + (k) < \Theta_\nu$;

b) $(j) + (k) < \Delta_\alpha$ for all α. This is not possible because $\Delta_1 = 0$;

or c) $(j) + (k) < \Theta_\alpha$ if and only if $(j) + (k) < \Delta_\alpha$ for all α. This is not possible because $\Delta_1 = 0$. Thus Ψ_1 is empty.

2. $(j) + (k) < \Psi_2$ implies $s_{ij}s_{jk}s_{ki} = 0$ for $s_{ij}s_{jk}s_{ki} = 1$ implies that i, j, and k are either a clique or a subset of a clique (by the argument of theorem 3), but $(j) + (k) < \Psi_2$ implies j and k are not contained in any clique.

3. Ψ_3 gives that

$$s_{ii}^{(3)} = \sum_{(j)+(k)<\Psi_3} \sum s_{ij}s_{jk}s_{ki}$$

$$= \sum_{\nu=1}^{m} \left\{ \sum_{\substack{(j)+(k)<\Theta_\nu \\ (j)+(k)<^*\Delta_\nu}} \sum s_{ij}s_{jk}s_{ki} \right\} .$$

We observe that: $\Omega_1[j, k \mid (j) + (k) < \Theta_\nu] = \Omega_2[j, k \mid (j) + (k) < \Theta_\nu, (j) + (k) < \Delta_\nu] + \Omega_3[j, k \mid (j) + (k) < \Theta_\nu, (j) + (k) <^* \Delta_\nu]$ and since $\Omega_2 \cdot \Omega_3 = 0$, it follows that $\sum_{\Omega_1} = \sum_{\Omega_2} + \sum_{\Omega_3}$ or $\sum_{\Omega_3} = \sum_{\Omega_1} - \sum_{\Omega_2}$. Ω_1 is the set of all ordered pairs $(j) + (k) < \Theta_\nu$. If $i \neq j \neq k \neq i$, then $s_{ij} = s_{jk} = s_{ki} = 1$, otherwise one of the $s_{pq} = 0$. Since every Θ_ν contains t_ν elements, there are ${}_{t_\nu - 1}P_2$ ordered pairs satisfying these conditions. Thus:

$$\sum_{\Omega_1} = {}_{t_\nu-1}P_2 = (t_\nu - 2)(t_\nu - 1) .$$

Similarly

$$\sum_{\Omega_2} = \begin{cases} (d_\nu - 2)(d_\nu - 1), & \nu > 1 \\ 0 & \nu = 1 \quad \text{since} \quad \Delta_1 = 0 . \end{cases}$$

Combining these,

$$\sum_{\Omega_3} = \begin{cases} (t_\nu - 2)(t_\nu - 1) - (d_\nu - 2)(d_\nu - 1), & \nu > 1 \\ (t_\nu - 2)(t_\nu - 1) & , \nu = 1 . \end{cases}$$

Summing over ν gives

$$s_{ii}^{(3)} = \sum_{\nu=2}^{m} \{(t_\nu - 2)(t_\nu - 1) - (d_\nu - 2)(d_\nu - 1)\}$$

$$+ (t_1 - 2)(t_1 - 1)$$

$$= \sum_{\nu=1}^{m} \{(t_\nu - 2)(t_\nu - 1) - (d_\nu - 2)(d_\nu - 1)\} + 2.$$

Since the entries $s_{ij}^{(3)}$ are uniquely determined from the entries of S by the laws of matrix multiplication, all valid methods of calculating $s_{ii}^{(3)}$ will give the same result. Specifically, in the above formula the numbering of the cliques is immaterial.

Similar formulas to that just deduced may be given for the off-diagonal terms of S^3, but they are considerably more complex, and, to date, they have not been found useful in applications.

5.08. Theorem 5: If 1) Θ is a set of t members with $t \geqslant 3$, 2) $s_{ii}^{(3)} = (t-2)(t-1)$ for i contained in Θ, and 3) $s_{jj}^{(3)} = 0$ for j contained in Θ', then Θ is a clique of t members.
Proof: There are two cases:

1. $i <=> j$ for all $i, j \,\varepsilon\, \Theta$, then Θ is a clique by definition 4 (§5.03) and theorem 3 (§5.06), and it has t members by part 1 of the hypothesis.

2. There exist p and $q \,\varepsilon\, \Theta$ such that p and q are not symmetric. Then by definition

$$S_{ii}^{(3)} = \sum_{(j)+(k)<\Theta} \sum s_{ij}s_{jk}s_{ki}$$

$$+ \sum_{(j)+(k)<^*\Theta} \sum s_{ij}s_{jk}s_{ki}.$$

If $s_{ij}s_{jk}s_{ki} = 1$, the elements i, j, and k are a clique or a subset of a clique and thus by hypothesis (3) and theorem 3 (§5.06) they are all contained in Θ; therefore the second sum $= 0$. Introduce in Ξ sufficient relationships $p => q$ to make Θ a clique Φ of t members. Since $s_{ij} \geqslant 0$ for all i and j, the introduction of these $s_{pq} = 1$ must increase the sum by 2 or more, for at least two additional 3-chains are introduced (i,p,q,i and i,q,p,i); hence by theorem 4 (§5.07)

$$S_{ii}^{(3)} = \sum_{(j)+(k)<\Phi} \sum s_{ij}s_{jk}s_{ki} - 2 = (t-2)(t-1) - 2$$

$$< (t-2)(t-1),$$

which is contrary to hypothesis (2). Therefore Θ is a clique of t members.

5.09. Redundancies:

By definition 3 (§5.03) an n-chain is redundant if and only if it contains at least one redundant pair (k,m), where a redundant pair defines two members of the n-chain γ_k and γ_m with $\gamma_k = \gamma_m$ and $k < m$. If these ordered subscript pairs (k,m) and the end point pair (i,j) (the latter not necessarily a redundant pair) are considered as sets, then five classes of mutually exclusive redundant n-chains may be defined which include all redundant n-chains:

1. The A_n class: There exists at least one redundant pair (k,m) and it has the property:

$$(k,m) \cdot (i,j) = 0 .$$

2. The B_n class: There exists one and only one redundant pair (k,m) and it has the property:

$$(k,m) \cdot (i,j) = i .$$

3. The C_n class: There exists one and only one redundant pair (k,m) and it has the property:

$$(k,m) \cdot (i,j) = j .$$

4. The D_n class: There exist two and only two redundant pairs (k,m) and (p,q) and they have the properties:

$$(k,m) \cdot (i,j) = i$$
$$(p,q) \cdot (i,j) = j .$$

5. The E_n class: There exists one and only one redundant pair (k,m) and it has the property:

$$(k,m) \cdot (i,j) = (i,j) .$$

If there are t n-chains $i \overset{(n)}{=>} j$ of the class A_n from i to j, then define $a_{ij}^{(n)} = t$. From these numbers the matrix $A^{(n)} = [a_{ij}^{(n)}]$ is formed. This is the matrix of redundant n-chains of the class A_n. If $R^{(n)}$ is the matrix of redundant n-chains it follows, if analogous definitions are made for matrices of the other four classes, that

$$R^{(n)} = A^{(n)} + B^{(n)} + C^{(n)} + D^{(n)} + E^{(n)} .$$

It follows directly from the definitions and the limitations on n that

$$R^{(1)} = 0$$
$$R^{(2)} = [\delta_{ij} x_{ij}{}^{(2)}] = E^{(2)}$$
$$A^{(3)} = 0\,.$$

It will now be proved that $D^{(3)} = S$. By the definition of the class D_3, there exist two and only two redundant pairs (k,m) and (p,q), and they have the properties:

$$(k,m)\cdot(i,j) = i$$
$$(p,q)\cdot(i,j) = j\,.$$

These pairs may define in total either three or four members of the 3-chain (three members when $m = p$, but no fewer for if $k = p$ and $m = q$ then $(k,m)\cdot(i,j) = (i,j)$, which is contrary to the definition of D_3). Suppose $m = p$, then either $i = \gamma_2 = j$ or $i = \gamma_3 = j$, which is impossible for $i \neq> i$ by assumption. Thus $m \neq p$. With four members there are two possibilities for a redundant 3-chain: either $i = \gamma_2$, $\gamma_3 = j$ or $i = \gamma_3$, $\gamma_2 = j$. The former is impossible by the previous argument; thus the only 3-chains of the class D_3 are of the form

$$i\,,\gamma_2\,,\gamma_3\,,j \equiv i\,,j\,,i\,,j\,;$$

that is,

$$d_{ij}{}^{(3)} \begin{array}{ll} = 1 & \text{if} \quad i <=> j \\ = 0 & \text{otherwise.} \end{array}$$

Therefore, by the definition of S, we have $D^{(3)} = S$.

If the matrices of redundancies up to and including $R^{(n-2)}$ are known, then we can find $A^{(n)}$ by $A^{(n)} = XR^{(n-2)}X$.

Proof: By the definition of the class A_n, a redundant n-chain of this class has the form

$$i \equiv \gamma_1\,,\gamma_2\,,\overset{(a)}{\text{——}}\,,\gamma_k\,,\overset{(b)}{\text{——}}\,,\gamma_m\,,\overset{(c)}{\text{——}}\,,\gamma_n\,,\gamma_{n+1} \equiv j\,,$$

where $a + b + c + 5 = n$, $k < m$, and $\gamma_k = \gamma_m$.

It follows from the definition that $p \equiv \gamma_2 \overset{(n-2)}{=\!=\!>} \gamma_n \equiv q$ is a redundant n—2 chain, and each such distinct n—2 chain determines no more than one distinct redundant n-chain from i to j. Thus the number of redundant n-chains of type A_n from i to j is the sum over all combinations $p \equiv \gamma_2$ and $q \equiv \gamma_n$ for the number of redundant n—2 chains from p to q, that is,

$$a_{ij}{}^{(n)} = \sum_{(p)+(q)<\equiv}\sum x_{ip} r_{pq}{}^{(n-2)} x_{qj}$$

or

$$A^{(n)} = XR^{(n-2)}X\,.$$

If the matrix $[e_{ij}^{(n)}]$ is defined as

$$[e_{ij}^{(n)}] = XR^{(n-2)}X + D^{(n)}$$

then the relations

$$A^{(n)} + B^{(n)} + D^{(n)} = R^{(n-1)}X$$
$$A^{(n)} + C^{(n)} + D^{(n)} = XR^{(n-1)}$$
$$E^{(n)} = [\delta_{ij}(x_{ij}^{(n)} - e_{ij}^{(n)})]$$

follow through an enumeration of cases and by using similar patterns of proof to that just given.

These various relations permit the specific conclusions:

$$R^{(2)} = [\delta_{ij}x_{ij}^{(2)}] = E^{(2)}$$
$$R^{(3)} = XR^{(2)} + R^{(2)}X + E^{(3)} - S$$

and the general result

$$R^{(n)} = XR^{(n-1)} + R^{(n-1)}X - XR^{(n-2)}X + E^{(n)} - D^{(n)} .$$

This latter expression is not useful in its present form because $D^{(n)}$ has not been expressed in terms of the matrices of redundancies up to and including $R^{(n-1)}$. This problem of the determination of the matrix of redundant n-chains is left as an unsolved problem of both theoretical and practical interest.

5.10. Uniqueness:

In certain applications it is desirable to know whether a power of a matrix uniquely determines the matrix. This is not true in general, for Sylvester's theorem gives a multiplicity of n^{th} roots of a matrix. The matrices being considered are rather specialized, however, and it is possible that some degree of uniqueness may exist.

The following two theorems indicate certain sufficient conditions for uniqueness. Since these theorems do not utilize completely the special characteristics of the matrices in this study, it is probable that more appropriate theorems can be proved.

5.11. Theorem 6: If p and q are positive integers, if two integers a and b can be found such that $ap - bq = 1$, and if X is a non-singular matrix, then the powers X^p and X^q uniquely determine X.
Proof: Suppose that there exist two non-singular matrices X and Y such that $X^p = Y^p$ and $X^q = Y^q$. Then $X^{ap} = Y^{ap}$ and $X^{bq} = Y^{bq}$. Now, form $X^{bq}Y = Y^{bq}Y = Y^{bq+1} = Y^{ap}$, since $ap - bq = 1$. Similarly $X^{bq}X = X^{bq+1} = X^{ap}$. But since $X^{ap} = Y^{ap}$ it follows that $X^{bq}X = X^{bq}Y$.

Since X is non-singular, $|X^{bq}| \neq 0$, and thus there exists a unique inverse of X^{bq}, X^{-bq}, such that $X^{-bq}X^{bq} = I$; therefore $X = Y$.

5.12. Theorem 7: If n is a positive odd integer and S a real symmetric matrix, then there is one and only one real symmetric n^{th} root of S.

Proof: 1. There is one such n^{th} root.

Since S is real and symmetric there exists a real orthogonal matrix P such that $P'SP = D$ (P' is the transpose of P) is diagonal with real entries d_{ii} which are the characteristic roots of S.* Assume P is so chosen that $d_{11} \leqslant d_{22} \leqslant \cdots \leqslant d_{mm}$. Let B be the diagonal matrix of the real n^{th} roots of the elements of D, i.e., $b_{ii} = \text{real } (d_{ii})^{1/n}$, so

$$B^n = D. \tag{1}$$

Define $R = PBP'$. Then $R^n = S$, for

$$R^n = (PBP')^n = PB^nP' = PDP' = S.$$

Since B is real and diagonal and P is real and orthogonal, R is real and symmetric.

2. There is only one such n^{th} root.

Suppose there exists a real symmetric matrix R_1 not equal to R such that $R_1{}^n = S$. Then there exists an orthogonal matrix Q such that $Q'R_1Q = T$ is diagonal in the characteristic roots of R_1, and ordered as before. Consider the n^{th} power of T:

$$\begin{aligned} T^n &= (Q'R_1Q)^n = Q'R_1{}^nQ = Q'SQ \\ &= Q'PDP'Q = (P'Q)'D(P'Q) \\ T^n &= U'DU, \end{aligned} \tag{2}$$

where U is the orthogonal matrix $P'Q$. Since $U' = U^{-1}$, T^n and D are similar, and hence have the same characteristic roots.† Because they are diagonal in the characteristic roots, ordered in the same way, they are equal:

$$D = T^n. \tag{3}$$

Substituting (3) in (2)

$$D = U'DU$$

or

$$UD = DU.$$

*MacDuffee, C. C. Vectors and matrices. Ithaca, N. Y.: The Mathematical Association of America, 1943, pp. 166-170.
†Ibid., p. 113.

By definition of matrix multiplication this means

$$\sum_j u_{ij} d_{jk} = \sum_j d_{ij} u_{jk} .$$

Since D is diagonal, this reduces to

$$u_{ik} d_{kk} = d_{ii} u_{ik}$$

or

$$u_{ik} (d_{kk} - d_{ii}) = 0 . \tag{4}$$

Since the d_{kk} are real and n is odd, equation (4) implies

$$u_{ik} [(d_{kk})^{1/n} - (d_{ii})^{1/n}] = 0$$

where the $(d_{kk})^{1/n}$ are real. Thus by the definition of B ,

$$UB = BU$$

or

$$B = U'BU . \tag{5}$$

By (1) and (3)

$$T^n = D = B^n ,$$

but by construction T and B are both real diagonal matrices and n is odd, so this implies

$$T = B .$$

This substituted in (5) gives

$$T = U'BU = Q'PBP'Q$$

or

$$QTQ' = PBP' .$$

But $QTQ' = R_1$ and $PBP' = R$ by definition; therefore

$$R_1 = R .$$

6. *Acknowledgement*

We wish to acknowledge our indebtedness to Dr. Leon Festinger, Assistant Professor of Psychology, Massachusetts Institute of Technology, for his kindness in directing this research to useful ends, encouraging the application of this method to practical problems, and providing many constructive criticisms of the work.

PSYCHOMETRIKA—VOL. 14, NO. 2
JUNE, 1949

A NOTE ON THE ESTIMATION OF TEST RELIABILITY BY THE KUDER-RICHARDSON FORMULA (20)*

LEDYARD R TUCKER
EDUCATIONAL TESTING SERVICE

The Kuder-Richardson formula (20) is rewritten to be identical with the simplest formula, (21), except for the addition of a term involving the standard deviation, σ_p, of the item p's. If σ_p can be estimated, a rapid and superior estimate of test reliability is possible in contrast to the simpler formula (21) used when the number of items and mean and standard deviation of test scores are known.

Quick estimates of test reliability are frequently desired and the Kuder-Richardson formula (21) is often used. This formula, however, sometimes seriously underestimates the reliability of a test. The Kuder-Richardson formula (20) yields a much better estimate, but requires results from an item analysis of the test. Formula (20) here is rewritten to involve only the standard deviation, σ_p, of item p's in addition to the number of items and the mean and standard deviation of test scores.

Following the Kuder-Richardson notation, formulas (20) and (21) are:

$$r_{tt}=\left(\frac{n}{n-1}\right)\left(\frac{\sigma_t^2-n\,\overline{pq}}{\sigma_t^2}\right); \tag{20}$$

$$r_{tt}=\left(\frac{n}{n-1}\right)\left(\frac{\sigma_t^2-n\,\bar{p}\,\bar{q}}{\sigma_t^2}\right). \tag{21}$$

where r_{tt} is the test reliability; σ_t^2 is the variance of scores on the test; n is the number of items; p_i is the proportion of candidates giving the correct answer to item i; q_i is $(1-p_i)$; $\bar{p}$ is the mean p_i; $\bar{q}$, $(1-\bar{p})$, is the mean q_i; and $\overline{pq}$ is the mean p_iq_i. Equation (22) gives the relation of $\bar{p}$ to the mean score, M_t, on the test:

$$\bar{p}=\frac{M_t}{n}. \tag{22}$$

*Kuder, G. F. and Richardson, M. W. The theory of the estimation of test reliability. *Psychometrika.* 1937, 2, 151-160.

(Equation (22) is the last one given by Kuder and Richardson. The formulas in this note will be numbered from (23) on for simplicity.) Formula (21) is derived from formula (20) by assuming that all items are of equal difficulty. Our approach is to rewrite formula (20) so that it involves n, σ_t, M_t, and σ_p. Other more realistic assumptions of σ_p may then be made.

A formula for σ_p^2 is:

$$\sigma_p^2 = \frac{1}{n}\sum_{i=1}^{n} p_i^2 - \bar{p}^2. \tag{23}$$

But:

$$\begin{aligned}\overline{pq} &= \frac{1}{n}\sum_{i=1}^{n} p_i q_i ,\\ &= \frac{1}{n}\sum_{i=1}^{n} p_i\,(1-p_i)\,,\end{aligned} \tag{24}$$

$$= \bar{p} - \frac{1}{n}\sum_{i=1}^{n} p_i^2. \tag{25}$$

Substituting equation (23) in equation (25):

$$\begin{aligned}\overline{pq} &= \bar{p} - \bar{p}^2 - \sigma_p^2\\ &= \bar{p}\,(1-\bar{p}) - \sigma_p^2\\ &= \bar{p}\,\bar{q} - \sigma_p^2 .\end{aligned} \tag{26}$$

Substituting equation (26) in formula (20):

$$r_{tt} = \left(\frac{n}{n-1}\right)\left(\frac{\sigma_t^2 - n\,\bar{p}\,\bar{q} + n\sigma_p^2}{\sigma_t^2}\right), \tag{27}$$

Or, using equation (22):

$$r_{tt} = \left(\frac{n}{n-1}\right)\left(\frac{\sigma_t^2 - M_t + \dfrac{M_t^2}{n} + n\sigma_p^2}{\sigma_t^2}\right). \tag{28}$$

Equation (27) is quite similar to formula (21), since only one term, $n\sigma_p^2$, is added in the numerator.

It is immediately apparent that formula (21) holds whenever σ_p is zero. However, better estimates of σ_p can be made if item difficulties are known. Ideally, σ_p is to be computed from an item analysis of the test in its final form, but estimates can be made from analy-

ses of the items in experimental forms of the test or in other tests. It might even be possible to guess a practicable value of σ_p from editorial judgment of the test items. The results for formula (21) are seldom more than 10 per cent less than those for formula (20). Consequently, the term involving σ_p contributes less than 10 per cent to the estimate of the test reliability. Thus even an error of 20 per cent in σ_p^2 will result in only a 2 per cent error in the estimated test reliability.

Results for two tests are summarized in the following table.

					Reliability		
							Odd-Even
Type of Test	n	M_t	σ_t	σ_p	*Formula 21*	*Formula 27*	*(Corrected)*
Verbal	140	63.5	21.5	.052	.93	.95	.95
Mathematical Aptitude	55	19.5	7.9	.057	.81	.86	.87

In these cases n was found by a count of the number of items, M_t and σ_t were computed from a frequency distribution of total scores, and σ_p was obtained from an item analysis of the test in its final form. Two useful types of limits for σ_p^2 are that: a) for a rectilinear distribution of item p's from .00 to 1.00, σ_p equals .083, and b) for a normal distribution with .00 at $-3\sigma_p$ and 1.00 at $+3\sigma_p$, σ_p^2 equals. 028.

PSYCHOMETRIKA—VOL. 14, NO. 2
JUNE, 1949

APPLICATION OF THE CONCEPT OF SIMPLE STRUCTURE TO ALEXANDER'S DATA*

MARIANO YELA
THE UNIVERSITY OF CHICAGO

A battery of 20 tests originally analyzed by Alexander (1) was reworked according to the principle of simple structure. His results were sustained in general. Both analyses yielded five factors in the first-order domain. Of these, three factors in the re-analysis (v, X and F) have almost exactly the same loadings as the corresponding factors in the original work, and were interpreted in the same way. The loading pattern of a fourth factor, Z, left uninterpreted in the original study, happened tó be more clear in the re-analysis, and an interpretation was attempted. It appears to be a factor of perceptual synthesis, and seems to play an important role in intellectual processes. A fifth factor, not present in Alexander's results, appeared in the new analysis: the reasoning factor, involved in inductive and deductive thinking. All four cognitive factors are related to a general factor that can be thought of as representing abstraction and eduction of relations and correlates, these processes being, therefore, the essential feature underlying intellectual behavior, at least in that sector surveyed by the tests of the present battery.

Among the papers dealing with the existence of a general factor, Alexander's work (1) is one of the most interesting and carefully done. It is, for instance, the first attempt to reconcile Thurstone's methods with Spearman's theory.

The clarity and thoroughness of his testing and exposition, together with the importance of his results, on the one hand, and the fact that we believe better techniques of rotation have been made available since his work, on the other, moved us to try a new factorial study of his data.

The Work of Alexander

The paper of Alexander presents the analysis of four batteries of verbal, performance, mechanical, and perceptual tests, given to four different experimental populations. An adequate discussion of the testing procedure is given in his paper, so that we do not need to be concerned here with this topic.

*The analysis of the data was done in the Psychometric Laboratory of The University of Chicago, under a fellowship from The University of Madrid (Junta de Relaciones Culturales). The author wishes to express his gratitude to Professor L. L. Thurstone, whose advice as a scientist and kindness as a friend, have been the principal stimulus for this work.

The four correlation matrices were factored by the centroid method. Subsequent orthogonal rotations led Alexander to final structures which he interpreted, arriving at five independent factors: a general intellective factor g, a verbal factor v, a performance factor F, a character factor X, which he named "persistence" or "will to succeed," and a factor Z the nature of which did not seem to be clearly indicated in the study.

Of the four groups studied in the original work we will be concerned here only with Group 3.†

The battery given to this group consisted of twenty tests, as follows:

1 Achievement in Shop Work (Engineering, etc.)
2 Achievement in Mechanical Drawing
3 Achievement in Mathematics
4 Achievement in Science
5 Achievement in English
6 Verbal Test 1 (Terman Group, 1, 2, 3, 4)
7 Verbal Test 2 (Terman Group, 6, 7, 8, 9)
8 Verbal Test 3 (Otis Self-Administering Test)
9 Verbal Test 4 (Otis Group, 1, 2, 3, 4, 7, 8)
10 Number-Verbal Test 1 (Terman Group, 5 and 10)
11 Number-Verbal Test 2 (Otis Group, 5 and 6)
12 Cox Test E3
13 Cox Test C
14 Cox Test D
15 The Passalong Test
16 Kohs' Block Design Test
17 The Cube Construction Test
18 Spearman's Test 1 (Form Series Test)
19 Spearman's Test 2 (Dot Pattern Test)
20 Spearman's Test 3 (Form Analogies Test)

Discussion of these tests and references to the literature can be found in the original paper.

Alexander's method led him to the rotated orthogonal matrix given in Table 2. The factors represented in this table were defined by Alexander as follows:

†Groups 1, 2, and 4 were analyzed by Alexander into three factors: g, v, and F. We factored and rotated to simple structure these correlation matrices and found three factors in each case: v, F, and a factor which could be called induction or reasoning. These results are not reported in full because they are in agreement with the results obtained in the analysis of Group 3, which will be fully reported, and because, having only three primary factors, we cannot have any assurance in the study of the second-order domain.

A general factor g. As this factor was common to all tests in the battery and as its corresponding axis was located through one of Spearman's tests of g, it was identified with the general factor of Spearman, and was defined as a general intellective factor.

A verbal factor v. This factor was loaded in all verbal tests and absent from all others. Factor v was defined as that factor independent of g which is found in verbal tests of intelligence.

A performance factor F. This factor was most conspicuous in the Cube Construction Test, with a loading of .70, and had high saturations on Cox's tests and lower saturations on Block Design, Passalong, and Shop Work. It was defined as that factor independent of g which is found in performance tests of intelligence.

A temperament or character factor X, present in all school achievement tests and absent from all others. It was defined as the factor present in school achievement and independent of g and v, and was named "persistence" or "will to succeed."

A factor Z, uninterpreted but proposed as related to school achievement.

Discussion of Alexander's Method

From the viewpoint of multiple-factor analysis as it is generally understood today in America, one of the important merits of Alexander's work is the fact that he realized, as early as 1935, the importance of rotation in the factorial procedure.

As a result of factoring the correlation matrix one has an orthogonal reference frame which is arbitrary in the sense that the location of the frame in the test space depends on the mathematical method of factoring adopted. Alexander states that this statistical frame has to be rotated to a psychological frame before the interpretation can be undertaken. By this he means that the axes should be placed in the location of the greatest psychological significance. And, he adds, a vital point is that the axes should be orthogonal or uncorrelated, because we are looking for independent factors. How do we know, however, that a certain frame has the greatest psychological significance?

The criterion adopted by Alexander is to pass the axis through a chosen vector. Thus he needs, of course, another criterion for the

selection of the test to be used as pivot. He goes then to previous findings in factorial studies. He chose one of the tests offered by Spearman and Stephenson as a good measure of g, and passed the first axis through it. Another test already identified as a good measure of v was chosen as the pivot for the verbal factor, and the subsequent axes were drawn through the clusters of residuals left after the effect of the previously identified factors had been ruled out of the tests.

Now, we cannot avoid the suspicion that if we use our previous hypothesis or the result of previous experiments as a criterion to locate the frame, we run the risk of forcing the results to be in accordance with our expectations. If the results turn out to be psychologically meaningful, this can be due to the correctness of our hypothesis or to the fact that they agree with what we previously thought of as a meaningful psychological theory because we have used this psychological theory as a directing criterion in our work.

What we need in factorial analysis is a criterion independent of our particular theory and the findings of previous factorial studies. We do not need to be concerned beforehand with the problem of which tests would give more meaning to the frame. We need first of all a criterion as independent as possible of any specific theory and as dependent as possible on the properties of the test configuration itself. The ideal would be to *discover* the structure that is demanded by the configuration, instead of *imposing* an assumed meaningful frame upon the configuration.

There is an infinite number of structures corresponding to an infinite number of reference frames. Mathematically any frame and consequently any structure is as good as any other. But perhaps there is a particular reference frame which is especially connected with the configuration, such as to be strongly demanded by the configuration itself. This is the assumption underlying the concept of simple structure (Thurstone, 2, 319-346). This assumption only goes as far as to gamble on the existence of planes well defined and overdetermined in the test configuration and by the test configuration itself. In Thurstone's own words: "If an analysis is to be made by the principles of simple structure, then the investigator gambles that the complexity of each test or measurement is less than the complexity r of the battery as a whole." (2, 320). Furthermore, it is believed that the appearance of simple structure is not a matter of chance (2, 328). Both statements spring from the basic assumption: "In the interpretation of mind we assume that mental phenomena can be identified in terms of distinguishable functions, which do not all participate equally in everything that mind does." (2, 57).

In any particular investigation it is, then, a matter of fact whether the simple structure is present or not, and whether the discovered simple structure can or cannot be meaningfully interpreted, or, if interpreted, whether it agrees or not with our previous hypothesis or with the findings of previous studies. In this light we do not need to impose the condition that the axes should be orthogonal. Whether they are orthogonal or oblique depends on the configuration. With this we do not lose the independence of our factors. It is necessary to distinguish between linear and statistical independence. We want to know the dimensionality of the psychological structure. Whether the parameters which represent these dimensions are statistically correlated or not is a fact to be found out rather than to be postulated.

All this does not mean that any other method of rotation is necessarily wrong. The literature shows how different methods have arrived at comparable results. Alexander's assumptions seemed to us quite reasonable, so that we believed from the beginning that our result would not differ greatly from his, except perhaps in the general factor. What we mean is that, in our opinion, the method of simple structure is, among those now available, the most rigorous and objective to check a previous hypothesis and the most flexible to explore a new domain.

Results of This Study

The centroid matrix F, given in Alexander's paper (1, 109; our Table 1), was rotated to simple structure. Table 3 shows the obtained oblique factor matrix V. To facilitate comparisons we have placed Alexander's final matrix and ours together (Tables 2 and 3), as well as the corresponding factor patterns (Tables 5 and 6). Table 4 gives the transformation matrix Λ.

It can be seen that our loadings are almost exactly the same as those found by Alexander for the four factors present in both matrices. To make the comparison easier we have reversed axis Z in Alexander's structure.

Within the limits imposed by the high complexity of some of the tests, the structure here presented satisfies the standard requirements for simple structure (2, 335 ff.).

Factors v, F, and X have the same pattern and lead to the same interpretation in both studies. Factor Z shows a more clear pattern in our study as a result of the application of the concept of simple structure.

The interpretation of the factors is as follows:

FACTOR Z

Principal saturations on factor Z

Code number	Name	Factor loading
16	Kohs' Block Design	.47
19	Dot Pattern Test	.47
18	Form Series Test	.39
15	Passalong Test	.27

All these tests require the subject to form or complete a configuration. This factor seems to represent speed of closure of a pattern to be formed following some formal instructions and against conflicting or irrelevant elements.

The subject will excel in this task if he can hold the given structure as a group of elements organized into a pattern and at the same time can reproduce it quickly (Tests 16, 19, and 15), or is able to perceive the figure that completes the unfinished configuration (Test 18). At the beginning of the task the elements integrate themselves into changing configurations that interfere with the completion of the correct pattern. In all cases the final structure is arrived at by quickly rejecting the patterns that do not lead to the correct configuration and by the ability to synthesize the units given into a meaningful whole.

This factor is similar to some others revealed in the recent literature. Thurstone (3) found a factor of perceptual closure (factor *A*), and another of flexibility in manipulating several irrelevant or conflicting "Gestalts" (factor *E*). In both factors the Block Design Test had significant saturations. In a battery of perceptual tests Bechtoldt (4) found a factor *G*, speed of closure of a visual configuration similar to our synthetic factor; he also found another factor *Y* similar to Thurstone's *E*, which he defines as a facility in organizing several simultaneous or successive configurations into a larger pattern under the distraction of further activity. Other studies such as those of Meili (5) and Rimoldi (6) present similar factors. Rimoldi reports a factor *B* as involved in the perception of relations in space necessary for the construction of a whole, and a factor *C* as a facility in solving the conflict between two or more configurations. These factors are related to those called "globalization" and "plasticité" by Meili. Factor *Z* is probably a composite of these two types of synthetic factors, which are not necessarily opposite to each other, as Meili pointed out (5, 43). Furthermore, in these studies there were found some interesting relations between this synthetic factor or factors and reasoning. Thurstone (3) reports a correlation of .39 between a com-

posite test of factor A and another of factor R (reasoning). Also he found that the reasoning tests had a saturation of .42 on factor E.

Bechtoldt reports a closure factor of second order, tentatively interpreted as a facility in forming conceptual structures, in which two primaries had loadings: the perceptual closure factor G and a factor of speed of ideational closure with verbal materials V. In our study the primary Z shows a correlation of .59 with R (reasoning), as shown in the correlation matrix R (Table 7). Both primaries Z and R have the highest saturations in the second-order general factor (Table 10). These results present some additional evidence to the findings of recent research in pointing out the importance of a closure or synthetic perceptual factor in the performance of tests of intelligence.

Several theories and clinical observations have in the past suggested the existence of such a trait. We should remember in this connection the theory of the synthetic sense of Aristotle and the scholastic psychologists (7), and the contention of the Gestalt theory that the ability to form "Gestalts" and the freedom from "Gestaltbindung" are fundamental features in productive thinking.

Examination of the other tests in the battery shows that this perceptual closure is not specially required in their solution. A perceptual synthesis of the material may be necessary in all thinking, but only in those tests demanding special synthetic ability would it be responsible for individual differences. There are, however, two tests that at first sight seem to depend directly on this ability. These are the Cube Construction and the Form Analogies Tests.

In the Cube Construction Test (cf. 1, 42 and 154 f.) the subject is required to reproduce a model by manipulating items in a way similar to that demanded by Tests 15 (Passalong) and 16 (Kohs' Blocks). A closer study of the tests and the subjects' performances would show, however, that in Tests 15 and 16 the subject works with the elements as pieces related to one another forming a whole. He is looking for the best way to arrive at a configuration. He perceives not independent elements, but partial configurations, and these he perceives related to the remaining items until a total closure is attained. This is clearly so in Kohs' Blocks and to a lesser extent in the Passalong Test. Accordingly, Test 16 has a loading of .47 and Test 15 a loading of .26 in the factor.

In Test 17 (Cube Construction), on the contrary, the items of the model are partially hidden. To see the model as a configuration and to fit the items together, the subject has to visualize in his mind the different possible positions of the blocks. He is likely to work with each item separately, especially in the difficult models, and place

it without a clear idea of the total structure. If this description is right we would expect this test to have a high saturation in some spatial factor. Actually we have a space factor in the analysis, and this test has a loading of .70 on this factor. Tests 15 and 16 require evidently this sort of visualizing of the figures as they change in space, and accordingly they also have some loading in this factor (.25 for Test 15, and .29 for Test 16).

The scoring of these tests supports this interpretation. Tests 15 and 16 are scored as a whole. The task is right or wrong and the subject gets a score for the time consumed in the successful performances. In Test 17 the score is a function of time and success per unit; i.e., the task is broken up in little tasks and each block correctly placed is considered as a success regardless of the position of the others.

The Form Analogies Test (number 20) is the only one among the remaining tests in which some element has to be picked up out of some irrelevant ones to closure an incomplete perceptual configuration. This test has a loading of .19 on the synthetic factor. This loading is probably insignificant, so that our results do not warrant any further discussion of it.

One question regarding this factor remains. Test 1 (Shop Work Achievement) has a negative loading on this factor (—.38). It is not clear how the facility to integrate diverse elements into a configuration would harm shop work performance. This test also has a high negative saturation in the corresponding factor of Alexander. All other tests in our configuration are included within reasonable limits in a positive region. And then, Test 1 stands out alone. This may be due to some computational error. Or maybe it is due to some actual characteristic of the variable. That we cannot decide from the information given in the original paper.

FACTOR R

Principal saturations on factor R

Code number	Name	Factor loading
10	Terman Group Test (Subtests 5 and 10)	.46
11	Otis Group Test (Subtests 5 and 6)	.52
20	Form Analogies	.37

Test 10 is composed of arithmetic problems and geometric figures (a circle, a triangle, and a rectangle with numbers and directions). Test 11 consists of arithmetic problems and number series. Test 20 requires the subject to choose among several forms the one related to a second as a third is to a fourth. The common feature involved in

the performance of these tests is the ability to find out a general principle by analysis of the given elements or the application of a rule to solve a problem or to identify an element. Some tests, as the mathematical problems, involve generally both processes, which can be called inductive and deductive thinking. Accordingly we may call this a reasoning factor. The complexity of Tests 10 and 11 may be responsible for the fact that we find a single factor for inductive and deductive reasoning, since some indications of two separate factors have been reported in the literature. Or it may be that both processes can be a function of a single factor.

Other tests in the battery have lower loadings in this factor. Test 8 is the only verbal test with an appreciable loading (.26) in this factor. Inspection of the test shows that it is the only verbal test having among its items a number of arithmetic problems and the same subtest of geometric figures as Test 10.

Form Series has a loading of .21 on the reasoning factor. This test had also, it will be remembered, a loading of .39 on the synthetic factor. Examination of the test shows that it can be solved in two ways or that two factors are required in its performance. The fourth form which continues the series initiated by the first three forms may be found out by closure of the unfinished configuration, perceiving the lacking element as connected with the others, or by analytically reasoning out the principle connecting the items. These processes have been recognized by Spearman to be present in the solution of the Raven matrices (8), and have been called by him the synthetic and the analytic approach respectively. Spearman states that the analytic procedure, not the synthetic, tends to load noegenetic processes with g. The analytic procedure would tend to load the performance with R, in our factor pattern. The synthetic procedure would tend to load the performance with Z. It is likely that both processes are used by most subjects in the task.

Tests 12 and 13 have saturations of .24 and .21 in this factor. They are the Cox Tests E3 and C. They are proposed as measures of mechanical ability. They require the subject to find out solutions to mechanical problems presented by drawings. They require the subject to put into practice the mechanical principle involved in each problem together with the ability to visualize the position of the different pieces as they move in space. This is consistent with the fact that both tests have also a loading of .40 on the space factor.

Test 1 has a saturation of .35 on this factor. It has also a loading of .36 on the space factor. If we assume that Shop Work Achievement, together with the Cox Test, is a measure of mechanical abil-

ity, this ability would be a function of the space and reasoning factors, since these tests have saturations in both.

FACTOR F

Principal saturations on factor F

Code number	Name	Factor loading
17	Cube Construction	.70
12	Cox Test E3	.40
13	Cox Test C	.40
14	Cox Test D	.27
1	Shop Work Achievement	.36
16	Block Design	.29
15	Passalong	.25
2	Mechanical Drawing Achievement	.20

This factor has the same loadings as factor *F* of Alexander. He interpreted it as a performance factor. Considering more recent factorial literature, we can conclude that this factor is the space factor *S*, as indeed has been already suggested (cf. Carroll, **9**, 308).

FACTOR V

Principal saturations on factor V

Code number	Name	Factor loading
6	Verbal Test 1	.70
5	English Achievement	.70
7	Verbal Test 2	.66
9	Verbal Test 4	.59
4	Achievement in Science	.59
8	Verbal Test 3	.44
3	Achievement in Mathematics	.45

This factor is readily identified with the verbal comprehension factor, and Alexander's interpretation is sustained. Perhaps it should be pointed out that since Alexander's work a verbal fluency factor *W* (Thurstone, **10**) and an ideational fluency factor *F* (Taylor, **11**) have been discriminated as different from the verbal comprehension factor *V*.

FACTOR X

In this factor only the achievement tests have saturations (.76, .65, .55, .49, .42). Our results agree fully with those of Alexander so that we do not need to add anything to the discussion and interpretation of this factor presented in pp. 125 f. of his paper.

The Second-Order Domain

Factorization of matrix R (Table 7), resulted in the factor matrix F_2 (Table 8).

Since the reference axes m_2 are arbitrary and no simple structure was over-determined by the five primaries, we rotated the axes so that the five primaries would be included in a positive quadrant, on the assumption that the second-order factors would not be negatively correlated with the first-order primaries. The oblique axes are shown in matrix V_2 (Table 10). Table 9 gives the transformation matrix for the second-order factors.

A factor appears in the second-order domain common to all cognitive primaries. The loadings are as follows: reasoning .76, perceptual synthesis .71, verbal .59, space .50, character factor X .00. This factor seems to be independent of X, in which we see a corroboration of Alexander's belief that factor X does not belong to the field of cognition. The structure of the second-order domain is not stable enough to attach much confidence to these results, but the correspondence with Alexander's findings is worth noticing. We do not believe that from this study nor from the total factorial research done so far a final answer can be presented as to the nature of the general intellective factor. If something is common to all cognitive functions, the perceptual synthesis, the verbal comprehension, the manipulation of space images, and the inductive and deductive thinking, to consider only the field covered by Alexander's battery, this would be the capacity of the subject to understand his task and his working, i.e., to see the meaning of the words, percepts, spatial relations, and logical relations, and to see the relation of these meanings to one another and to the solution of the problem. This general feature in all intellectual tasks has been named by Spearman "abstraction and noegenesis" (8). This explanation seems logical. The validity of this explanation is difficult to ascertain from the factorial studies alone. We believe, however, that the factorial way of attacking the problem is the analysis of batteries with cognitive and non-cognitive tests, so constructed that a second-order domain can be expected to have a simple structure as overdetermined as those so far found in the first-order domain. If a general intellective factor is present, as a great deal of evidence indicates, its nature will be better understood when we know which factors are related to it, which are independent, and how much the related factors are loaded in the general factor. By the application of the rotational principle of simple structure we, then, were able to get the same information as Alexander, plus some interesting new items, as reported above. This we take as supporting our belief that Thurstone's principle of rotation affords a more fruit-

ful approach to the problem and, contrary to some opinions, does not preclude the finding of a general factor.

TABLE 1

Alexander's Centroid Matrix F*

	I	*II*	*III*	*IV*	*V*	h^2
1	403	187	350	064	515	589
2	602	276	575	—057	063	776
3	660	353	244	—186	062	658
4	620	505	291	—040	—008	726
5	581	574	—023	—056	—083	678
6	715	334	—422	189	—190	873
7	700	221	—416	173	—293	828
8	820	052	—326	012	—032	784
9	767	184	—318	—046	—182	759
10	785	015	—172	—273	157	745
11	706	—098	—148	—282	221	658
12	651	—205	—075	194	080	515
13	538	—316	—177	243	022	480
14	531	—215	054	069	001	336
15	458	—213	067	128	—175	307
16	540	—458	141	084	—257	594
17	431	—206	094	565	093	565
18	455	—387	089	—307	—098	469
19	414	—335	177	—186	—238	406
20	568	—282	—061	—264	059	479

*Decimal points have been omitted from this and subsequent tables.

TABLE 2
Alexander's Rotated
Orthogonal Matrix

	Z	*X*	*V*	*F*	*g*
1	–447	503	011	278	243
2	000	766	173	179	358
3	–118	563	351	–020	452
4	–095	633	477	051	292
5	–152	394	656	–082	259
6	–095	–065	822	122	402
7	063	–124	779	128	431
8	–114	–055	548	128	673
9	000	000	646	000	585
10	–219	098	276	–060	730
11	–232	065	153	–026	759
12	000	–029	211	398	558
13	032	–205	164	408	496
14	063	060	091	277	493
15	235	030	135	281	393
16	401	–004	–003	339	564
17	000	000	141	704	221
18	219	043	–138	–027	632
19	354	098	–076	050	512
20	000	000	000	000	692

TABLE 3
Our Oblique Matrix *V*

	Z	*X*	*V*	*F*	*R*
1	–385	489	–027	358	349
2	069	760	313	204	–022
3	000	546	448	009	112
4	–034	650	574	070	–058
5	–076	420	701	–073	–049
6	–040	–008	700	100	–002
7	095	–067	663	082	–057
8	022	–047	435	126	259
9	107	016	587	–015	104
10	–012	061	267	–015	462
11	–022	005	115	012	521
12	023	–017	079	399	242
13	063	–187	–001	395	210
14	143	056	042	272	152
15	269	045	104	250	–034
16	473	–004	–014	293	–023
17	–071	061	–041	696	026
18	393	–018	–072	–040	211
19	474	066	009	015	006
20	186	–057	–006	010	370

TABLE 4
Λ

	Z	*X*	*V*	*F*	*R*
I	122	237	396	270	251
II	–393	525	739	–265	–265
III	247	817	–167	248	–296
IV	–265	–020	007	837	–328
V	–836	020	–518	308	820

TABLE 5
Alexander's Factor Pattern

	Z	X	V	F	g
16	40			34	56
19	35				51
18	22				63
15	24			28	39
1	–45	50		28	24
2		77			36
3		56	35		45
4		63	48		29
5		39	66		26
6			82		40
7			78		43
8			55		67
9			65		58
17				70	22
12			21	40	56
13		–20		41	50
14				28	49
10	–22		28		73
11	–23				76
20					69

TABLE 6
Our Factor Pattern

	Z	X	V	F	R
16	47			29	
19	47				
18	39				21
15	27			25	
1	–38	49		36	35
2		76	31	20	
3		55	45		
4		65	57		
5		42	70		
6			70		
7			66		
8			44		26
9			59		
17				70	
12				40	24
13				40	21
14				27	
10			27		46
11					52
20					37

TABLE 7
Correlations between
Primaries, R

	Z	X	V	F	R
Z	1.000				
V	.228	–.258	1.000		
X	.083	1.000			
F	.346	–.178	.392	1.000	
R	.589	.168	.461	.258	1.000

TABLE 8
Orthogonal
Factor
Matrix F_2

	I_2	II_2
Z	608	438
X	-190	558
V	687	-200
F	563	-120
R	651	440

TABLE 9
Transfor-
mation
Matrix Λ_2

	G	W
I_2	953	216
II_2	305	977

Table 10
Factor
Matrix V_2

	G	W
Z	713	559
X	-011	504
V	594	-047
F	500	004
R	755	570

REFERENCES

1. Alexander, W. P. Intelligence, concrete and abstract. *Brit. J. Psychol., Monograph Supplements*, 1935, 6, No. 19.
2. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. Chicago Press, 1947.
3. Thurstone, L. L. A factorial study of perception. Chicago: Univ. Chicago Press, 1944.
4. Bechtoldt, H. P. Factorial study of perceptual speed. Unpublished Ph.D. dissertation. Department of Psychology, University of Chicago, 1947.
5. Meili, Richard. L'analyse de l'intelligence. *Archives de Psychologie*, 1946, 31, No. 121, 1-64.
6. Rimoldi, H. J. A. Study of some factors related to intelligence. *Psychometrika*, 1948, 13, 27-47.
7. Moore, T. V. The synthetic sense and intelligence. *Psychol. Rev.*, 1938, 45, 219-227.
8. Spearman, C. Theory of general factor. *Brit. J. Psychol.*, 1946, 36, 117-131.
9. Carroll, J. B. The factorial representation of mental ability and academic achievement. *Educ. psychol. Meas.*, 1943, 3, 307-332.
10. Thurstone, L. L. Primary mental abilities. *Psychometric Monograph No. 1*. Chicago: Univ. Chicago Press, 1938.
11. Taylor, C. W. A factorial study of fluency in writing. Unpublished Ph.D. dissertation: Department of Psychology, University of Chicago, 1946.

PSYCHOMETRIKA—VOL. 14, NO. 2
JUNE, 1949

DEVELOPMENT OF A METHOD FOR INCREASING THE UTILITY OF MULTIPLE CORRELATIONS BY CONSIDERING BOTH TESTING TIME AND TEST VALIDITY*

W. F. LONG AND IRVING W. BURR
PURDUE UNIVERSITY

A modification of the Wherry-Doolittle test selection method is presented by which tests are included in a multiple correlation (obtained for a given battery of tests) in the sequence in which the rate of return in validity per unit of testing time is greatest, rather than in the order of the size of their contribution to the multiple correlation. It is proposed that the modified method can be utilized profitably when there are economic or practical limits on the time available for test administration.

Introduction

In situations where selection of personnel is based, at least in part, on test results, multiple correlation coefficients are being used more and more extensively. Such a multiple correlation is a measure of the relationship between a criterion and a battery of tests. After a multiple correlation is obtained, a multiple regression equation is calculated from which the criterion can be predicted with the highest precision of which the given battery of tests is capable when using a first-degree equation.

Possibly the best available method for general use for obtaining such multiple correlations is the Wherry-Doolittle test selection method.† It makes possible the selection from a number of tests of the battery or team of tests which gives the maximum possible multiple correlation with a minimum number of tests, and at the same time saves an appreciable amount of statistical work in comparison with other methods. The obtained multiple correlation is the best estimate of the *R* for the population from which the sample was drawn, or in other words, it is corrected for the tendency of correlations obtained

*The major portion of this article is based upon a thesis by W. F. Long directed by Dr. Joseph Tiffin with the counsel of Dr. Irving W. Burr. This thesis was submitted in partial fulfillment of the requirements for the degree of Master of Science in Psychology, Purdue University, June, 1947.

†See Garrett, Henry E. Statistics in psychology and education. New York: Longmans, Green & Co., 1947, pp. 435-451 or Stead, W. H., Shartle, D. L. and associates. Occupational counseling techniques. New York: American Book Co., 1940, Appendices 5 and 6.

from samples to be larger than the correlation existing in the total population.

One of the convenient features of the Wherry-Doolittle test selection method is the fact that tests are selected for inclusion in the multiple correlation in the order of their contribution to the correlation. For example, after the first test is selected, the test with the highest residual validity is selected as the second test to be added to the battery. Next, the test with the then highest residual validity is selected as the third test, and so on, until the multiple R ceases to increase in size by an amount greater than the chance error introduced by the tests included. The term "residual validity" is applied to the remaining validity that a test has after the effect of the intercorrelations of that test with other tests is discounted as the battery is formed.

Purpose.—In the application of test selection batteries, other factors besides validity are of considerable importance, such as ease of administration, cost of tests, and testing time required. Of these factors, testing time required is especially important from an economic point of view.

Accordingly, it was considered desirable to develop a procedure in which testing time as well as validity of individual tests is considered when tests are evaluated for inclusion in a multiple correlation, and thus for eventual use in a selection test battery.

The procedure developed is a simple modification of the Wherry-Doolittle test selection method designed so that tests are included in the multiple correlation in the sequence in which the rate of return in validity per unit of testing time is largest, rather than in the order of their contribution to the multiple.

Both the Wherry-Doolittle method and the modified method have been applied to the same battery of fifteen tests in order to permit comparison of the two methods and to furnish an application of the latter method. A presentation of the development of the modified method and a detailed outline of the procedure for its use will be followed by a brief description of the demonstration battery, a description of the applications of the modified method, and a comparison of the results obtained.

Development of Modified Method

In order to determine which test will furnish the greatest return of validity per unit of testing time, the formula

$$r_{c(ax)} = \frac{a\bar{r}_{cx}}{\sqrt{a + (a^2 - a)r_{1I}}}, \tag{1}$$

which is based on the Spearman-Brown phophecy formula,* is used.† This formula must be modified so that predicted validities for different lengths of a test can be calculated when the given validity of a test is a residual validity and not the original validity of the test, as yet unaffected by its correlation with other tests. The formula then becomes

$$r_{c(ax)R} = \frac{a\sqrt{\dfrac{V_x^2}{Z_x}}}{\sqrt{a + (a^2 - a)r_{1IR}}}. \tag{2}$$

In these two formulas:

$r_{c(ax)}$ = validity of test X (correlation between test X and criterion) the length of which, in terms of time, has been multiplied by a factor a. It is important to note that in using a time value for a, it is assumed that the time for the test is changed in proportion to the change in test content, i.e., a longer test would include more items with the same ratio between time and number of problems.

$\bar{r}_{cx}$ = original validity of test X.

r_{1I} = original reliability of test X.

$r_{c(ax)R}$ = residual validity of test X, the length of which has been multiplied by a factor a.

$\dfrac{V_x^2}{Z_x}$ = square of the residual validity of test X. This value is what remains of $\bar{r}_{cx}^2$ (square of the validity of test X) after the effect of the intercorrelations of test X with the other tests is discounted as the battery is formed.

r_{1IR} = residual reliability of test X.

The value r_{1IR} can be estimated in terms of the original reliability of the test and the residual validity of the test. This can be calculated by use of the formula

$$r_{1IR} = \sqrt{\frac{r_{1I}^2 + \dfrac{V_x^2}{Z_x} - r_{cx}^2}{1 + \dfrac{V_x^2}{Z_x} - r_{cx}^2}}. \tag{3}$$

*Use of this formula assumes that the entire test is homogeneous.

†Peters, Charles C. and Van Voorhis, Walter R. Statistical procedures and their mathematical bases. New York: McGraw-Hill Book Co., 1940, p. 196.

The derivation of this formula is not complicated:

$\sqrt{r_{cx}^2 - \frac{V_x^2}{Z_x}}$ = the loss from the original validity of test X because of the intercorrelations of the test with tests already included in a battery.

r_{1I} = original reliability of test X.

r_{1I}^2 = proportion of variance in a second equivalent form of the test (for example) explained by the first form.

$(1 - r_{1I}^2)$ = $(\text{error})^2$ or r_{1I}^2 + $(\text{measurement error})^2 = 1^2$ (perfect reliability).

If much of the correlation of test X with the criterion is accounted for by other tests already included in a battery, the reliable part of test X which has not yet been included may not have as large a reliability coefficient as the total test had. Let it be assumed that the equivalent of several perfectly reliable items are taken from test X, leaving all the factors making for error still in the test. Now it can be determined how much the reliability of the original test could have been decreased.

Thus

$$r_{1IR}^2 = \frac{r_{1I}^2 - \left(r_{cx}^2 - \dfrac{V_x^2}{Z_x}\right)}{r_{1I}^2 - \left(r_{cx}^2 - \dfrac{V_x^2}{Z_x}\right) + 1 - r_{1I}^2}. \tag{4}$$

Then

$$r_{1IR} = \sqrt{\frac{r_{1I}^2 + \dfrac{V_x^2}{Z_x} - r_{cx}^2}{1 + \dfrac{V_x^2}{Z_x} - r_{cx}^2}}. \tag{5}$$

Reference to Table 4 will show that this correction is very small when applied for the demonstration battery, the largest being .0025. Since the corrected reliability values obtained by using this formula are so little different from the original values and do not enter into the actual calculation of the multiple correlation, it may well be that this step may prove to be unnecessary.

Modified Wherry-Doolittle Test Selection Method

A detailed procedure for use of the modified Wherry-Doolittle test selection method is outlined here in a general form so that it can be readily applied for the calculation of a multiple correlation for any specific battery of tests. It will be noted by those familiar with the Wherry-Doolittle procedure that the modified method differs from the original only in those steps involved in the determination of which test should be first included in the battery and the sequence in which the remaining tests should be added. These differences first appear in steps 4 and 10 of the modified procedure.

Given: a. Intercorrelations of all tests.*

b. Reliabilities of all tests.

c. Correlation of the tests with a criterion (validity).

d. Testing time for all tests.

These data can be handled most effectively if presented in a form similar to Table 1.

1. Prepare worksheets similar to Tables 2 and 3.
2. Enter in the V_1 row in Table 2 the validity coefficients of all tests *with signs reversed.*
3. Enter in the Z_1 row in Table 3 the number 1 for each test.
4. Select as the first test in the battery that test which will give the greatest rate of return in validity per unit of testing time. This is accomplished by completing the following steps:
   a. Prepare a worksheet similar to Table 5.
   b. Choose the test which has the largest quotient $\dfrac{V_1^2}{Z_1}$.
   c. Calculate time-adjusted validity values for this test and all tests requiring less testing time so as to permit comparison of these tests matched in regard to testing time with the shortest test, or at other strategic testing times. (Explained in detail on page 149). This is accomplished by application of the formula

$$r_{c(ax)R} = \frac{a\sqrt{\dfrac{V_x^2}{Z_x}}}{\sqrt{a + (a^2 - a)\, r_{1IR}}}, \qquad (2)$$

*Actually, only the intercorrelations of those tests included in the multiple correlation are needed. However, it is probably as economical in time and effort in the long run to calculate all intercorrelations at one time, unless the number of tests under consideration is large.

in which a is the new length of the test and r_{1IR} is the reliability of the test. The capital letter, R, generalizes the formula to make it applicable when $\frac{V_x^2}{Z_x}$ and r_{1IR} are residual values after the effect of inclusion of other tests in the battery is discounted.

In the worksheet just prepared:

(1) Enter in the series column, the series number.

(2) Enter in the test number column the test numbers of all tests necessary to be considered.

(3) Enter in the testing time column the testing times of all tests to be considered.

(4) Enter in Column A the quotient $\frac{V_1^2}{Z_1}$ of each test to be considered.

(5) Enter in Column C the multiplication factor necessary to make the testing time of each test equivalent to the time of the shortest test.

Then for each test in turn except the shortest:

(6) Enter in Column F the reliability of each test.

(7) Record in Column B the square root of the Column A entry.

(8) Record in Column D the product of the Column B and Column C entries.

(9) Record in Column E the difference between the square of the Column C entry and the Column C entry.

(10) Record in Column G the product of the Column E and Column F entries.

(11) Record in Column H the sum of the Column G and Column C entries.

(12) Record in Column J the square root of the Column H entry.

(13) Enter in Column K the quotient of the Column D entry divided by the Column J entry. This is the time-adjusted validity value (the predicted validity of the test in question for a different testing time). If the time-adjusted validity of the shortest test is equal to or greater than the time-adjusted validities of the other

tests considered, it should be selected for inclusion in the battery. If this is not true, all tests should be equated to make their testing times equal to the testing time of the second shortest test. If either the shortest or second shortest test then has the highest time-adjusted validity, that test should be selected for inclusion in the battery. If not, this process must be repeated in the indicated sequence until it is apparent which test will give the greatest rate of return in validity per unit of testing time. After a very few such calculations have been made, it is relatively easy to anticipate which test is most likely to have the highest time-adjusted validity value at a given time, and thus selection of the tests is not so complicated as it may seem. (Reference to the explanation of this step on page 149 as applied to the demonstration battery will facilitate its ready application).

5. Apply the Wherry shrinkage formula,

$$\bar{R}^2 = 1 - K^2\left(\frac{N-1}{N-M}\right)$$

in which $\bar{R}$ is the shrunken multiple correlation coefficient, $K^2 = 1 - \frac{V_1^2}{Z_1}$, N is the number of subjects in the sample, and M is the number of tests so far included in the battery.

This is accomplished by completing the following steps:

a. Prepare a worksheet similar to Table 6.
b. Enter in the test number column the number of the selected test.
c. Enter in Row O, Column C the number 1.
d. Enter in Row 1, Column B the quotient $\frac{V_1^2}{Z_1}$.
e. Record in Row 1, Column C the difference between the Row O, Column C entry and the Row 1, Column B entry.
f. Record in Column D the quotient $\frac{N-1}{N-M}$.
g. Record in Column E the product of the Column C and Column D entries.

h. Record in Column F the difference between 1 and the Column E entry.

i. Record in Column G the square root of Column F. This is the shrunken multiple correlation coefficient. Since the shrinkage factor, $\frac{N-1}{N-M}$, is unity when M, the number of tests in the battery, equals 1, the first $\bar{R}$ equals the coefficient of correlation between the selected test and the criterion.

6. Next, use the Doolittle method for solving normal equations.

This is accomplished by completing the following steps:

a. Prepare a worksheet similar to Table 7.

b. Leave the a_1 Row blank.

c. Enter in the b_1 Row the correlation coefficient of the first selected test with every other test, as well as with the criterion. The sign of the latter correlation coefficient must *always* be reversed in this table.

d. Record in the check sum column the algebraic sum of the entries.

e. Record in the c_1 Row the product of each b_1 entry and the negative reciprocal of the b_1 entry for the selected test. Formula: $c_1 = b_1$ (each test) $\times \frac{-1}{b_1}$ (selected test).

7. Draw a vertical line under the first selected test in Tables 2 and 3.

8. To each V_1 entry for the tests in Table 2, add algebraically the product of the b_1 entry in the criterion column and the c_1 entry for each of the other tests (from Table 7) to obtain the V_2 entries. Formula: $V_2 = V_1 + [b_1$ (criterion) $\times c_1$ (each test)$]$.

9. To each Z_1 entry for the tests in Table 2 add algebraically the product of the b_1 and c_1 entries of the corresponding tests (from Table 7) to obtain the Z_2 entries.
Formulas: $Z_2 = Z_1 + [b_1$ (each test) $\times c_1$ (same test)$]$

10. Select as the second test in the battery the test which will give the greatest rate of return in validity per unit of testing time. The selection of the second test is accomplished exactly as in Step 4 except that the residual reliability of

those tests entered in Table 5 must be determined. The $\frac{V_2^2}{Z_2}$ value of the test so selected is a measure of the amount which the second test contributes to the squared multiple correlation coefficient, $\bar{R}^2$.

The calculation of the residual reliabilities is accomplished by completing the following steps:

a. Prepare a worksheet similar to Table 4.
b. Enter in the series column the series number.
c. Enter in the test number column the test numbers of all tests to be considered.
d. Enter in Column A the original reliability of each test.
e. Enter in Column C the $\frac{V_2^2}{Z_2}$ values of each test. (The subscript always is the same as the series number).
f. Enter in Column E the original validity for each test.
g. Record in Column B the square of the Column A entry.
h. Record in Column D the sum of the Column B and Column C entries.
i. Record in Column F the square of the Column E entry.
j. Record in Column G the difference between the Column D and Column F entries.
k. Record in Column H the sum of the Column C entry and 1.
l. Record in Column J the difference between the Column H and Column F entries.
m. Record in Column K the quotient obtained by dividing the Column G entry by the Column J entry.
n. Record in Column L the square root of Column K. The values in Column L are the residual reliabilities of the tests to be used in Table 5.

11. Again apply the shrinkage formula as in Step 5, using Table 6.
 a. Enter in the test number column the test number of the second selected test.
 b. Enter in Row 2, Column B the quotient $\frac{V_2^2}{Z_2}$.

c. Record in Row 2, Column C, the difference between the Row 1, Column C entry and the Row 2, Column B entry.

d. Record in Column D the quotient $\dfrac{N-1}{N-M}$.

e. Record in Column E the product of the Column C and Column D entries.

f. Record in Column F the difference between 1 and the Column E entry.

g. Record in Column G the square root of the Column F entry. This value for $\bar{R}$ is the new shrunken multiple correlation coefficient. If it is smaller than the preceding $\bar{R}$, the second test has added more chance error than actual validity. In this event, work should be stopped and only the first test should be used to predict the criterion. If the $\bar{R}$ is larger than the preceding one, the addition of tests to the battery must be continued.

12. Continue the Doolittle procedure, using Table 7.

a. Enter in the a_2 Row of Table 7 the correlation coefficients of the second selected test with every other test, as well as with the criterion. Remember to reverse the sign of the correlation of the second selected test with the criterion.

b. Record in the check sum column the algebraic sum of the a_2 entries.

c. Draw a vertical line through the b_2 and c_2 Rows for the first selected test.

d. Record in the b_2 Row the sum of each a_2 entry and the product of the b_1 entry of the same test and the c_1 entry for the second selected test. Formula: $b_2 = a_2 + [b_1$ (each test) $\times\ c_1$ (second selected test)]. The entries in the Criterion and Check Sum Columns are determined by the same method.

e. There are three checks in the b_2 Row:

(1) The entry for the second selected test should equal the Z_2 entry for the same test in Table 3.

(2) The entry in the criterion column should equal the V_2 entry of the second selected test in Table 2.

(3) The entry in the check sum column should equal the sum of all the other entries in the b_2 row.

f. Record in the c_2 Row the product of each b_2 entry and the negative reciprocal of the b_2 entry for the second selected test. Formula: $c_2 = b_2$ (each test) $\times \frac{-1}{b_2}$ (second selected test).

g. There are three checks in the c_2 Rows:
   (1) The c_2 Row entry of the second selected test should be -1.
   (2) The c_2 Row entry in the check sum column should equal the sum of all the other c_2 entries.
   (3) The product of the b_2 and c_2 entries of the criterion column should equal the $\frac{V_2^2}{Z_2}$ value in Table 6, Column B, Row 2. (Disregard sign).

13. Draw a vertical line under the second selected test in Tables 2 and 3.

14. The V_3 entries are calculated exactly as in Step 8 except that all subscripts should be increased by 1 to apply at this point. Formula: $V_3 = V_2 + [b_2(\text{criterion}) \times c_2 \text{ (each test)}]$

15. The Z_3 entries are calculated exactly as in Step 9 except that all subscripts should be increased by 1 to apply at this point. Formula: $Z_3 = Z_2 + [b_2(\text{each test}) \times c_2 \text{ (same test)}]$

16. Select as the third test in the battery, the test which will give the greatest return in validity per unit of testing time. This is accomplished in the same manner as in Steps 10 and 4.

17. Apply the shrinkage formula as in Step 5 using Table 6. If the $\bar{R}$ is larger than the preceding one, a fourth test must be considered for inclusion in the battery.

18. Continue the Doolittle procedure using Table 7.
   a. Enter in the a_3 Row of Table 7 the correlation coefficients of the third selected test with every other test as well as with the criterion, the latter with reversed sign.
   b. Record in the check sum column the algebraic sum of the a_3 entries.
   c. Draw a vertical line through the b_3 and c_3 entries of previously selected tests.
   d. The formula for the b_3 entries is: $b_3 = a_3 + [b_1$ (each test)

$\times$ c_1 (third selected test)] + [b_2 (each test) $\times$ c_2 (third selected test)].

e. The formula for the c_3 entries is: $c_3 = b_3$ (each test) $\times$ $\frac{-1}{b_3}$ (third selected test).

f. The same checks apply here as given in parts e. and g. of Step 12.

19. The values for V_4 and Z_4 are determined as in Steps 14 and 15 with the required change in subscript. Formulas: $V_4 = V_3$ + [b_3 (criterion) $\times$ c_3 (each test)]; $Z_4 = Z_3$ + [b_3 (each test) $\times$ c_3 (same test)].

20. Repeat the necessary calculations in steps 4 through 10 until an $\bar{R}$ smaller than its preceding one is obtained, until all the tests have been included in the battery, or until the potential gain from adding further tests seems uneconomical.

Test Selection by Modified Method

Content of Demonstration Battery.—The battery of tests used here for demonstration purposes includes 15 tests from among a group of 46 that had been administered, at various stages of their training, to 407 A.A.F. rated bombardiers who were candidates for radar observer training. The 15 tests were found to be correlated with the criterion at least at the five per cent level of significance.

The criterion was a composite of course grades which were based entirely on standardized test and performance check scores. Although no method was available for directly determining the reliability of the composite course grades, an approximate reliability was determined using the split-half method. The composite score was divided into two parts which were made as nearly alike as possible as to content. For a sample of 278 men from one training station, the correlation between the two "halves" of the course grade was .27, which became .43 when corrected for double length.

This approximated value for the reliability of the criterion does not enter into the multiple correlation calculations but it does serve as an approximate indication of the upper limit that the multiple R can be expected to reach.* The size of the obtained multiple R would therefore tend to indicate that the battery would do as good a prediction job as could be expected under the circumstances.

*Theoretically, the square root of the obtained reliability represents the upper limit of the coefficient of validity for a test. See Lindquist, E. F. A first course in statistics. New York: Houghton-Mifflin Co., 1942, p. 224.

The fact that eleven of the 15 tests are included in the multiple R before it starts to shrink can be at least partially accounted for by several reasons, among them being the relatively large size of the sample and generally low intercorrelations between the tests. The relatively low correlations of the individual tests with the criterion can be explained, to a certain degree at least, by the narrowed range of talent in the sample caused by the men being subjected to at least three selection procedures for which corrective data were not available, and by the relatively low reliability of the criterion.

Application of Modified Method

Referring to Table 2, Series 1, Test 8 has the largest $\frac{V_1^2}{Z_1}$, which is .0324. This test requires ten minutes of testing time. There are three tests, Numbers 10, 13, and 14, which require less testing time; hence they should be checked to determine whether one of them will furnish a higher rate of return in validity per unit of testing time than Test 8. Reference to Table 5, Series 1, wherein the calculated time-adjusted values of these tests appear, demonstrates that Test 8 for 7.5 minutes retains the highest adjusted value (in comparison with Tests 10 and 14) and is thus selected as the first test for inclusion in the multiple R. Test 14 was given a time-adjusted value for 7.5 minutes because it was anticipated that this test would not furnish as great a rate of return in validity per unit of testing time as Test 10. However, a check must be made, and 7.5 minutes was a convenient point, since the time-adjusted values of Tests 8 and 10 were calculated at that time. It was not necessary to calculate a time-adjusted value of Test 13 because its $\frac{V_1^2}{Z_1}$ value at five minutes testing time is considerably smaller than that of Test 14 at just 5.5 minutes and therefore would remain smaller when adjusted. As mentioned in Step 13 of the procedure previously presented, after a very few such calculations it is quite simple to anticipate in most cases which test and testing time will be selected, although of course this must be checked mathematically.

In Series 2, Table 2, Test 10, requiring 7.5 minutes testing time, has the largest $\frac{V_2^2}{Z_2}$ value. Tests 13 and 14, which require 5 and 5.5 minutes of testing time, respectively, have values that will not match the value of Test 10 at 7.5 minutes, as was indicated by the time-adjusted values determined in Series 1. Therefore Test 10 is the second test included in the battery.

In Series 3, Test 4 has the largest $\frac{V_3^2}{Z_3}$ value and has a shorter testing time than Tests 1 and 5, which are the only ones having values appreciably rivaling the value of Test 4, and therefore Test 4 is the third test selected for inclusion in the multiple.

In Series 4, Tests 1 and 5 have by far the largest $\frac{V_4^2}{Z_4}$ values. When a time-adjusted value is calculated for the latter test at 18 minutes for comparison with the former, Test 1 still has the largest time-adjusted value.

In Series 5, Tests 5 and 15 are equated for testing time, which shows that Test 15 has the largest time-adjusted value and is therefore the fifth test included in the muliple R.

In Series 6, Tests 5 and 13 are equated for testing time, with Test 5 again having the smaller time-adjusted value, indicating that Test 13 should next be included in the battery.

In Series 7, Test 5 is equated in time value to Test 14. At this point it has the largest time-adjusted value and therefore is the seventh test selected.

In Series 8, Test 15 has the largest $\frac{V_8^2}{Z_8}$ value with the shortest testing time, so it is included next in the multiple R without question.

In Series 9, Test 7, a 35-minute test, is equated with Test 2 at 20 minutes but retains the highest time-adjusted value and is thus the ninth test selected for inclusion in the battery.

In Series 10, Test 2 is compared with Test 6 and again has the smaller time-adjusted value, indicating that Test 6 should next be selected.

In Series 11 and 12 there is no question concerning which test should be added to the multiple.

Comparison of Results from the Two Methods

The results of the application of the regular Wherry-Doolittle test selection method to the demonstration battery are presented in Table 8. A comparison of the results obtained from the applications of the two methods is presented in Figure 1. In this figure cumulative testing time is plotted on the base line and the multiple R on the vertical axis. The first divergence of results from the two methods occurs when the fourth test is added to the battery.

Using the Wherry-Doolittle method, when Test 5 is added the multiple correlation becomes .283, requiring 63.5 minutes of testing time. With the modified method, Test 1 is added as the fourth test,

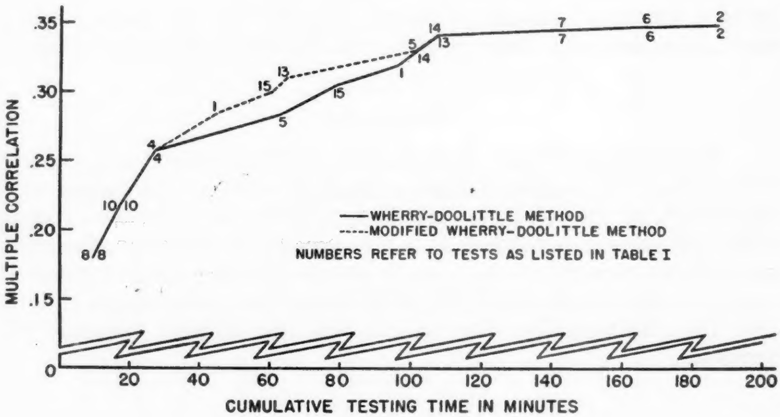

FIGURE 1
Comparison of Results from the Application of Original and Modified Wherry-Doolittle Method to a Demonstration Battery

making the multiple R exactly the same, but only 45.5 minutes of testing time is required. This difference of 18 minutes of testing time would be of considerable importance in an industrial testing situation. Adding a fifth test would produce a multiple $\bar{R}$ of .299, with just 60.5 minutes of elapsed testing time using the modified method. Again with the modified method, a multiple $\bar{R}$ of .309 is obtained using just two minutes more of testing time than is necessary to obtain an $\bar{R}$ of .283 using the Wherry-Doolittle method. This advantage of greater return in validity for elapsed testing time is maintained by using the modified method, as can be seen by reference to Figure 1, until the eighth test is added to make the obtained $\bar{R}$ .3395 by the Wherry-Doolittle method and .3393 by the modified method. The same three tests are subsequently added when using both procedures to bring the $\bar{R}$ to .3484 before it begins to shrink when another test is added. Since the $\bar{R}$ increases only from .339 to .348 with 80 minutes of additional testing time, it would probably not be practical to use the additional tests.

As a general practice in using a battery of tests selected on the basis of a multiple correlation, the number of tests that should be used for predictive purposes would depend considerably upon the significance of the difference between succeeding R's obtained as tests were added to the battery. Since we are interested here only in demonstrating the possibility of identifying test batteries which will re-

quire a minimum amount of testing time, it was not deemed necessary to present such differences for the battery used as an example.

Conclusions

The modified Wherry-Doolittle test selection method as presented will produce for a given battery of tests a multiple correlation with the maximum possible value requiring the minimum possible testing time. By using this method, it is also possible to determine the largest multiple correlation that can be obtained from a given battery of tests in a particular period of testing time. Such information is valuable for administrative reasons and has considerable economic importance.

TABLE 1
Intercorrelations of Fifteen Tests and a Criterion, Test Reliabilities, and Testing Times.* $N = 407$

Variable	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Criterion	Reliability	Testing Time
1	–.03	.04	–.06	.14	.15	.13	–.01	.12	.08	.17	.14	.04	.00	.07	.13	.89	18
2		.29	–.18	.01	.17	.19	.33	.47	.18	.02	.33	.26	.14	–.03	.11	.90	20
3			–.04	–.03	.23	.22	.37	.27	.19	.10	.35	.16	.24	.05	.12	.80	24
4				–.06	.14	.01	–.08	–.09	–.15	.02	–.02	–.22	–.21	–.11	.11	.88	10
5					.12	.15	–.11	.05	–.01	.32	–.04	–.03	–.19	–.04	.10	.87	36
6						.47	.24	.09	.11	.47	.24	.08	.07	.05	.10	.93	25
7							.33	.08	.13	.48	.27	–.04	–.12	.00	.14	.84	35
8								.33	.23	.12	.48	.14	.31	.18	.18	.68	10
9									.23	.06	.44	.20	.23	.17	.12	.90	15
10										.11	.27	.08	.26	.10	.17	.96	7.5
11											.17	–.06	–.03	.06	.10	.85	30
12												.18	.30	.18	.17	.81	15
13													.14	.03	.09	.97	5
14														.09	.13	.76	5.5
15															.14	.87	15

*Adapted from Table 11.13, Psychological research on radar observer training. Report 12, Army Air Forces Aviation Psychology Program Research Reports (Preliminary Draft).

TABLE 2
Worksheet for Selection of Tests for Inclusion in Multiple R, Part 1—Modified Wherry-Doolittle Method

	Test Number	1	2	3	4	5	6	7
	Testing Time, Minutes	18	20	24	10	36	25	35
$\frac{V_1^2}{Z_1}$								
	V_1	—.13	—.11	—.12	—.11	—.10	—.10	—.14
$\frac{V_2^2}{Z_2}$		.017373			.015575	.014528		
	V_2	—.1318	—.0506	—.0534	—.1244	—.1198	—.0568	—.0806
$\frac{V_3^2}{Z_3}$		.014651			.020762	.014026		
	V_3	—.1206	—.0365	—.0392	—.1423	—.1177	—.0494	—.0733
$\frac{V_4^2}{Z_4}$		.016494				.016510		.00505
	V_4	—.1278	—.0568	—.0386		—.1274	.—0251	—.0069
$\frac{V_5^2}{Z_5}$						.012563		
	V_5		—.0623	—.0341		—.1101	—.0050	—.0500
$\frac{V_6^2}{Z_6}$			.006452			.013487		
	V_6		—.0739	—.0368		—.1140	—.0040	—.0573
$\frac{V_7^2}{Z_7}$						.014248		
	V_7		—.0574	—.0274		—.1171	.0028	—.0651
$\frac{V_8^2}{Z_8}$								
	V_8		—.0521	—.0265			.0197	—.0450
$\frac{V_9^2}{Z_9}$			.003467					.005509
	V_9		—.0530	—.0154			.0235	—.0657
$\frac{V_{10}^2}{Z_{10}}$			.002620				.003990	
	V_{10}		—.0458	—.0056			.0529	
$\frac{V_{11}^2}{Z_{11}}$			.003154					
	V_{11}		—.0501	—.0104				
$\frac{V_{12}^2}{Z_{12}}$								
	V_{12}			—.0029				

TABLE 2 (Continued)
Worksheet for Selection of Tests for Inclusion in Multiple R, Part 1—Modified Wherry-Doolittle Method

Test Number	8	9	10	11	12	13	14	15
Testing Time, Minutes	10	15	7.5	30	15	5	5.5	15
$\frac{V_1^2}{Z_1}$	.0324		.0289			.0081	.0169	
V_1	—.18	—.12	—.17	—.10	—.17	—.09	—.13	—.14
$\frac{V_2^2}{Z_2}$			.017462			.004283	.006091	.011965
V_2		—.0606	—.1286	—.0784	—.0836	—.0648	—.0742	—.1076
$\frac{V_3^2}{Z_3}$								
V_3		—.0397		—.0672	—.0619	—.0583	—.0486	—.0996
$\frac{V_4^2}{Z_4}$		.002438			.004232	.008235	.006134	.013212
V_4		—.0459		—.0612	—.0560	—.0878	—.0718	—.1124
$\frac{V_5^2}{Z_5}$						.007597		.011445
V_5		—.0320		—.0398	—.0388	—.0843	—.0745	—.1044
$\frac{V_6^2}{Z_6}$						.007965	.006415	
V_6		—.0221		—.0369	—.0301	—.0863	—.0734	
$\frac{V_7^2}{Z_7}$							.005576	
V_7		—.0095		—.0440	—.0199		—.0683	
$\frac{V_8^2}{Z_8}$							.009643	
V_8		—.0004		—.0062	—.0196		—.0883	
$\frac{V_9^2}{Z_9}$								
V_9		.0101		—.0080	—.0063			
$\frac{V_{10}^2}{Z_{10}}$								
V_{10}		.0083		.0213	.0043			
$\frac{V_{11}^2}{Z_{11}}$								
V_{11}		.0102		.0056	.0031			
$\frac{V_{12}^2}{Z_{12}}$		.000143						
V_{12}		.0309		.0013	.0119			

TABLE 3
Worksheet for Selection of Tests for Inclusion in Multiple R, Part 2—Modified Wherry Doolittle Method

Test Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Z_1	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.	1.
Z_2	.9999	.8911	.8631	.9936	.9879	.9424	.8911		.8911	.9471	.9856	.7696	.9804	.9039	.9676
Z_3	.9927	.8797	.8515	.9753	.9877	.9392	.8880		.8660		.9784	.7427	.9780	.8663	.9640
Z_4	.9902	.8599	.8515		.9831	.9107	.8860		.8642		.9767	.7410	.9361	.8404	.9562
Z_5		.8581	.8503		.9649	.8861	.8687		.8525		.9488	.7231	.9354	.8399	.9523
Z_6		.8464	.8497		.9636	.8860	.8640		.8439		.9481	.7165	.9351	.8398	
Z_7		.8122	.8385		.9624	.8801	.8563		.8241		.9418	.7033		.8366	
Z_8		.8102	.8384			.8601	.8279		.8183		.8417	.7033		.8085	
Z_9		.8101				.8586	.7836		.8068						
Z_{10}		.8006				.7013			.8062						
Z_{11}		.7959							.8054						
Z_{12}									.6694						

TABLE 4
Worksheet for Selection of Tests for Inclusion in Multiple R, Part 3—Modified Wherry-Doolittle Method

		A	B	C	D	E	F	G	H	J	K	L
Series	Test Number	r_{1I}	r_{1I}^2	$\frac{V_x^2}{Z_x}$	$B+C$	r_{cx}	r_{cx}^2	$D-F$	$I+C$	$H-F$	$\frac{G}{J}$	$\sqrt{K}$
4	5	.87	.7569	.016510	.773410	.10	.0100	.763410	1.016510	1.006510	.758472	.8709
5	5	.87	.7569	.012563	.769463	.10	.0100	.759463	1.012563	1.002563	.757521	.8704
6	5	.87	.7569	.013487	.770387	.10	.0100	.760387	1.013487	1.003487	.757745	.8705
7	5	.87	.7569	.014248	.771148	.10	.0100	.761148	1.014248	1.004248	.757928	.8706
9	7	.84	.7056	.005509	.711109	.14	.0196	.691509	1.005509	.985909	.701392	.8375
10	6	.93	.8649	.003990	.868890	.10	.0100	.858890	1.003990	.993990	.864083	.9295

TABLE 5
Worksheet for Selection of Tests for Inclusion in Multiple R, Part 4—Modified Wherry-Doolittle Method

			A	B	C	D	E	F	G	H	J	K
Series	Test Number	Testing Time	$\frac{V_z^2}{Z_z}$	$\sqrt{A}$	Change in Length of Test	$B \times C$	C^2-C	r_{1IR}	$E \times F$	$G+C$	$\sqrt{H}$	$\frac{D}{J}$
1	8	10	.0324	.18	.75	.135	—.1875	.68	—.1275	.6225	.7890	.1711
	10	7.5	.0289	.17	1.							
	14	5.5	.0169	.13	1.3636	.177269	.4958	.76	.3692	1.7328	1.3163	.1347
4	1	18	.016494	.128429	1.							
	5	36	.016510	.128491	.5	.064246	—.25	.8709	—.2177	.2823	.5315	.1209
5	15	15	.011445	.106983	1.							
	5	36	.012563	.112085	.4167	.046695	—.2430	.8704	—.2115	.2051	.4529	.1031
6	13	5	.007965	.089245	1.							
	5	36	.013487	.116133	.1389	.016131	—.1196	.8705	—.1041	.0348	.1865	.0865
7	14	5.5	.005576	.074673	1.							
	5	36	.014248	.119365	.1528	.018239	—.1295	.8706	—.1127	.0401	.2002	.0911
9	2	20	.003467	.058885	1.							
	7	35	.005509	.074220	.5714	.042409	—.2449	.8377	—.2052	.3662	.6051	.07009
10	2	20	.002620	.051190	1.							
	6	25	.003990	.063169	.80	.050535	—.1600	.9295	—.1487	.6513	.8069	.0626

TABLE 6
Worksheet for Application of Wherry Shrinkage Formula—
Modified Wherry-Doolittle Method

Testing Time, Minutes	Cumulative Time, Minutes	Test Number	A M	B $\frac{V_x^2}{Z_x}$	C K^2 $(1-B)$	D $\frac{N-1}{N-M}$	E $\bar{K}^2$ $(C \times D)$	F $\bar{R}^2$ $(1-E)$	G $\bar{R}$ $(\sqrt{F})$
			0		1				
10	10	8	1	.0324	.9676	1.00	.9676	.0324	.1800
7.5	17.5	10	2	.01746	.95014	1.002469	.952486	.047514	.2180
10	27.5	4	3	.02076	.92938	1.004950	.933980	.066020	.2569
18	45.5	1	4	.01649	.91289	1.007444	.919686	.080314	.2834
15	60.5	15	5	.01144	.90145	1.009950	.910419	.089581	.2993
5	65.5	13	6	.00796	.89349	1.012469	.904631	.095369	.3088
36	101.5	5	7	.01425	.87924	1.015000	.892429	.107571	.3280
5.5	107.0	14	8	.00964	.86960	1.017544	.884856	.115144	.3393
35	142.0	7	9	.00551	.86409	1.020101	.881459	.118541	.3443
25	167.0	6	10	.00399	.86010	1.022670	.879598	.120402	.3470
20	187.0	2	11	.00315	.85695	1.025253	.878591	.121409	.3484
15		9	12	.00014	.85681	1.027848	.880670	.119330	.3454

TABLE 7
Worksheet for the Doolittle Solution of Normal Equations
Wherry-Doolittle Method

Test Number		1	2	3	4	5	6	7	8	9
	a_1	—	—	—	—	—	—	—	—	—
8	b_1	–.01	.33	.37	–.08	–.11	.24	.33	1.00	.33
	c_1	.01	–.33	–.37	.08	.11	–.24	–.33	–1.00	–.33
	a_2	.08	.18	.19	–.15	–.01	.11	.13	.23	.23
10	b_2	.0823	.1041	.1049	–.1316	.0153	.0548	.0541		.1541
	c_2	–.0869	–.1099	–.1108	.1389	–.0162	–.0579	–.0571		–.1627
	a_3	–.06	–.18	–.04	1.00	–.06	.14	.01	–.08	–.09
4	b_3	–.0494	–.1391	.0042	.9753	–.0667	.1668	.0439		–.0422
	c_3	.0507	.1426	–.0443	–1.00	.0684	–.1710	–.0450		.0433
	a_4	1.00	–.03	.04	–.06	.14	.15	.14	–.01	.12
1	b_4	.9902	–.0428	.0348		.1342	.1561	.1308		.1078
	c_4	–1.00	.0432	–.0351		–.1355	–.1576	–.1321		–.1089
	a_5	.07	–.03	.05	–.11	–.04	.05	.0001	.18	.17
15	b_5		–.1056	–.0249		–.0356	.0086	–.0669		.0905
	c_5		.1109	.0261		.0374	–.0090	.0703		–.0950
	a_6	.04	.26	.16	–.22	–.03	.08	–.04	.14	.20
13	b_6		.1789	.1024		–.0335	.0741	–.0847		.1360
	c_6		–.1913	–.1095		.0358	–.0792	.0906		–.1455
	a_7	.14	.01	–.03	–.06	1.00	.12	.15	–.11	.05
5	b_7		.0434	.0073		.9624	.1387	.1652		.0746
	c_7		–.0451	–.0076		–1.00	–.1441	–.1717		–.0775
	a_8	.0001	.14	.24	–.21	–.19	.07	–.12	.31	.23
14	b_8		–.0086	.1013			.0345	–.1892		.0963
	c_8		.0106	–.1253			–.0427	.2340		–.1191
	a_9	.13	.19	.22	.01	.15	.17	1.00	.33	.08
7	b_9		.0864	.1171			.3511	.7837		–.0216
	c_9		–.1102	–.1494			–.4480	–1.00		.0277
	a_{10}	.15	.17	.23	.14	.12	1.00	.47	.24	.09
6	b_{10}		.0575	.0632			.7014			–.0247
	c_{10}		–.0820	–.0901			–1.00			.0352
	a_{11}	–.03	1.00	.29	–.18	.01	.17	.19	.33	.47
2	b_{11}		.7958	.1188						.3289
	c_{11}		–1.00	–.1493						–.4133

TABLE 7 (Continued)
Worksheet for the Doolittle Solution of Normal Equations
Wherry-Doolittle Method

Test Number		10	11	12	13	14	15	—C	Check Sum
	a_1	—	—	—	—	—	—	—	—
8	b_1	.23	.12	.48	.14	.31	.18	–.18	3.68
	c_1	–.23	–.12	–.48	–.14	–.31	–.18	.18	–3.68
	a_2	1.00	.11	.27	.08	.26	.10	–.17	2.64
10	b_2	.9471	.0824	.1596	.0478	.1887	.0586	–.1286	1.7936
	c_2	–1.00	–.0870	–.1685	–.0505	–.1992	–.0619	.1358	–1.8938
	a_3	–.15	.02	–.02	–.22	–.21	–.11	–.11	–.1600
4	b_3		.0410	–.0406	–.2022	–.1590	–.0875	–.1423	.3835
	c_3		–.0420	–.0416	.2073	.1630	.0897	.1459	–.3932
	a_4	.08	.17	.14	.03	.0001	.07	–.13	1.8501
1	b_4		.1661	.1330	.0270	–.0213	.0623	–.1278	1.7505
	c_4		–.1677	–.1343	–.0273	.0215	–.0629	.1291	–1.7678
	a_5	.10	.06	.18	.03	.09	1.00	–.14	1.6601
15	b_5		.0265	.0790	–.0180	.0096	.9522	–.1044	.8110
	c_5		–.0278	–.0830	.0189	–.0101	–1.00	.1096	–.8517
	a_6	.08	–.06	.18	1.00	.14	.03	–.09	1.8700
13	b_6		–.0765	.1110	.9350	.0549		–.0863	1.3113
	c_6		.0818	–.1187	–1.00	–.0587		.0923	–1.4024
	a_7	–.01	.32	–.04	–.03	–.19	–.04	–.10	1.1800
5	b_7		.3104	.0019		–.1646		–.1171	1.4221
	c_7		–.3225	–.0020		.1710		.1217	1.4777
	a_8	.26	–.03	.30	.14	1.00	.09	–.13	2.1001
14	b_8		–.0161	.1219		.8080		–.0884	.8602
	c_8		.0199	–.1508		–1.00		.1093	–1.0640
	a_9	.13	.48	.27	–.04	–.12	.0001	–.14	3.1601
7	b_9		.3498	.1269				–.0657	1.7277
	c_9		–.4463	–.1619				.0838	–2.2045
	a_{10}	.11	.47	.24	.08	.07	.05	–.10	3.5300
6	b_{10}		.2083	.0158				.0530	1.0747
	c_{10}		–.2970	.0225				–.0756	1.5322
	a_{11}	.18	.02	.33	.26	.14	–.03	–.11	3.04
2	b_{11}		–.0679	.1390				–.0501	1.2643
	c_{11}		.0853	–.1747				.0630	–1.5887

TABLE 8
Worksheet for Application of Wherry Shrinkage Formula—
Wherry-Doolittle Method

			A	*B*	*C*	*D*	*E*	*F*	*G*
Testing Time, Minutes	Cumulative Time, Minutes	Test Number	M	$\frac{V_x^2}{Z_x}$	K^2 $(1—B)$	$\frac{N—1}{N—M}$	$\bar{K}^2$ $(C\times D)$	$\bar{R}^2$ $(1—E)$	$\bar{R}$ $(\sqrt{F})$
			0		1				
10	10	8	1	.0324	.9676	1.00	.9676	.0324	.18
7.5	17.5	10	2	.01746	.95014	1.002469	.952486	.047514	.2180
10	27.5	4	3	.02076	.92938	1.004950	.933980	.066020	.2569
36	63.5	5	4	.01651	.91287	1.007444	.919665	.080335	.2834
15	78.5	15	5	.01406	.89881	1.009950	.907753	.092247	.3037
18	96.5	1	6	.01084	.88797	1.012469	.899042	.100958	.3177
5.5	102	14	7	.01069	.87728	1.015000	.890439	.109561	.3310
5	107	13	8	.00771	.86957	1.017544	.884768	.115232	.3395
35	142	7	9	.00552	.86405	1.020101	.881418	.118582	.3444
25	167	6	10	.00399	.86010	1.022670	.879598	.120402	.3470
20	187	2	11	.00315	.85695	1.025253	.878591	.121409	.3484
		9	12	.00014	.85681	1.027848	.880670	.119330	.3454

PSYCHOMETRIKA—VOL. 14, NO. 2
JUNE, 1949

DON LEWIS. *Quantitative Methods in Psychology.* Ann Arbor: Edwards Brothers, 1948. Pp. v + 286.

The purpose of this volume is best characterized in the words of the author: "This book was prepared specifically as a text for an advanced graduate-level course in psychology, in the experimental area . . . . The contents of the present course, called Quantitative Methods in Psychology, are in general the same as the contents of this book. The course does not replace the usual course in statistics. It is given in addition to six other courses (three of them quite advanced) in the general area we call 'quantitative methods and statistics.' "

Since there are only ten chapters, the chapter headings are here listed to give the reader a quick view of the organization of the book.

Chapter 1—Variables, Constants, and Functional Relationships
Chapter 2—Fitting Curves to Empirical Data:
I. Linear Functions
Chapter 3—Logarithms
Chapter 4—Fitting Curves to Empirical Data:
II. Complex Functions
Chapter 5—Differentiation
Chapter 6—Integration
Chapter 7—The Normal Curve
Chapter 8—Distribution Functions
Chapter 9—Applications of Equations
Chapter 10—Goodness of Fit

The techniques discussed in the various chapters are thoroughly illustrated with concrete psychological data. Exercises appended to the chapters are well-chosen and great enough in number to give the student ample practice. The mathematical chapters should serve as adequate review for the student who has previously been through the calculus. We doubt that they would be adequate for the more typical psychology student who has not gone beyond college algebra, though this could hardly be considered the fault of their author.

Perhaps the most satisfactory chapters from our point of view were chapters 8 and 9. In the former the author describes the binomial and Poisson distributions, then discusses the relationships between the normal distribution and the distributions of t, chi square, and F in possibly the most meaningful manner available in psychological statistics. In the latter chapter the applications of equations both to physical science data and to psychological data are discussed. The seven pages devoted to Galileo and Newton, in this chapter, are quite illuminating. An expansion of this material would be much more useful to our students than courses in the philosophical history of psychology.

Since this is a preliminary edition, we feel free to suggest additions as well as corrections at this time. The reviewers, for instance, would have found a section on units of measurement in psychology very useful. The problem of units is certainly critical in the determination of functional relationships between variables. There has also been some degree of neglect of units of measurement in the preparation and discussion of the figures. There seem to be two rather different

approaches to the matter of graphing. One is aesthetic. The other is to attempt to portray essential relationships. In his frequent neglect of origins, and in his selection of units for the graphs, the author comes closer to the first of these two approaches.

It would help the reader, we believe, if the psychological implications of the various equations were sometimes discussed more fully. A learning curve equation on page 16, for example, has a negative value for the additive constant (y-intercept). It is surely of interest that the equation implies that learning starts below zero.

We also suggest that definitions could be emphasized to advantage by slight changes in typography and format. As it is, important definitions are frequently hidden in the text. We also found a few important terms introduced without definition; e.g., antilogs, on page 36.

There is a little confusion with regard to an application of chi square. On page 258 a paragraph is devoted to a discussion of a fundamental error made by certain other experimenters in the application of chi square to a series of 132 scores made by 33 subjects. Independence of separate scores had been implicitly assumed, but not tested. Previously, on page 162, the author made a similar mistake, involving 1600 scores from 320 animals, after having made a test which suggested, at the 5% level of confidence, that successive turns on the maze were not independent of each other. A further illustration on page 164 also involves multiple responses from individual animals, but a previous chi square test had suggested that the trials were statistically independent.

Other points, briefly noted, are as follows: (1) Some of the illustrations of constants (p. 1) are not particularly apt. The constancy of an individual's I.Q. is really quite a different sort of constancy than is the value of pi. (2) Definitions of dependent and independent variables (p.2) do not cover the many situations in which their designations are arbitrary. (3) Linear correlation is illustrated only as a measure of goodness of fit for scatters of means. The properties of r when used in this way are not distinguished from the properties of r when it is based upon scatters of individuals. For example, r in the first case is a function of the number of observations, whereas r in the second case is independent of N. (4) In the discussion of the normal curve as an outgrowth of the law of error (p. 145), the author does not make explicit all necessary assumptions regarding the nature of the errors. (5) A negative exponent (p. 186) should be changed to a positive when the accompanying expression is moved from the numerator to the denominator of the fraction. (6) The application of the F-test on page 193 involves doubling the p-value obtained from the table. (7) In the expression for the standard error of a difference between correlated means on page 194, the cross-product term is omitted. It is brought back implicitly in the next step by shifting to the distribution of differences. (8) The t-test suggested for use with independent distributions having unequal variances (p. 196) has certain merits, but it involves chance correlations between values paired at random. With samples of the size illustrated, chance correlations could markedly influence the final result. (9) The experimental advantages and disadvantages of the various t-theorems are largely neglected (pp. 189-196). (10) There are two significant omissions from the discussion of chi square. One is the failure to mention Yates' correction for continuity. The second is the failure to point out the suspicion with which one should view a very high chi square p-value, when testing goodness of fit. A chi square p greater than .95 will occur by chance no more often than a p less than .05, and should lead one to a careful check of computations and as-

sumptions. (11) A few typographical errors are evident, but are no more frequent than are expected in a lithoprinted publication of this nature.

The author has brought together a great deal of material not available elsewhere. Dr. Lewis has a much more comprehensive treatment than Guilford attempted in a somewhat comparable section of *Psychometric Methods*. The reviewers believe that this volume constitutes a major contribution to experimental-theoretical psychology. They further believe that advanced graduate students with these interests should be offered a course of the type described by this text.

Stanford University Lloyd G. Humphreys
Lyle V. Jones

EGON BRUNSWIK. *Systematic and representative design of psychological experiments.* Berkeley and Los Angeles: University of California Press, 1947. Pp. 60.

Experimental psychologists have long striven to emulate the rigorous control and isolation of variables in the manner of the physical sciences. Brunswik's monograph presents a thoughtful and challenging critique of psychological experimentation conceived in the image of physics and a call for a new type of experimental design. Although most of the illustrations are drawn from the area of perception, the thesis of the monograph has wide implications for the total field of quantitative research in psychology.

Representativeness is the concept which provides the key both to Brunswik's criticism of traditional experimental design and to his plea for a new methodology. The classical type of experiment usually employs only a few carefully chosen values of the stimulus variable: a set of sizes, weights, or brightnesses. The range of values of the variable which the organism encounters and responds to in its natural habitat is vastly larger and more diversified. From what Brunswik calls an *ecological* point of view, such a choice of variables fails to yield a representative sample of behavior. As a consequence, the results of many systematic classical experiments lack generality and cannot easily be translated into general laws about the organism's response to the variable.

Lack of representativeness, Brunswik further argues, also results from the manner in which the covariation of variables is handled in the classical experiment. Even though the observed behavior may be a joint function of several determinants, these determinants are not allowed to vary together in their natural (i.e., ecologically effective) manner. Instead, variables are "artificially tied," "artificially interlocked," or "artificially untied." The *artificial tying* of variables is illustrated by the conventional Galton-bar experiment, or for that matter, a large number of classical psychophysical situations. In such a situation, the physical size of the stimuli and their retinal projections are not allowed to vary independently; due to the artificial perfect correlation, it is impossible to assess the relative contribution of these two factors to perceived size. *Artificial interlocking* of variables is exemplified by the conventional size-constancy experiment. With the physical size of the stimulus constant, the magnitude of the retinal projection is allowed to vary over a few fixed values. Finally, variables are *artificially untied* when a large number of potential determinants is successfully held constant and only one factor is allowed to vary. When personality traits are judged, for example, subjects may be required to wear identical clothing and to assume the same position. The probably joint dependence of the judgment on

facial expression, posture, and clothing habits is arbitrarily destroyed. To the extent that variables cannot be held constant, they are frequently assigned to residual variance.

In Brunswik's analysis, the procedures of mental testing represent, as it were, the reverse side of the classical coin. Here individual differences are the focus of interest and quantitative treatment. But the stimulus (test) is held as constant as possible, and lack of stimulus variation precludes ecological generality of results. Neither the systematic, quasi-physical type of experimental design nor the differential statistics of mental testing thus meet Brunswik's requirement of representativeness.

These criticisms do not, as the author fully recognizes, apply with equal force to the total range of experimental designs used in psychological research today. Brunswik emphasizes the gradual but steady development of perceptual research away from single-variate investigations to the study of field dynamics or "multi-dimensional psychophysics." Multivariate experiments, however, are only a first step toward truly representative design. To achieve representativeness, it is necessary to apply the full freedom (and restrictions) of sampling to stimulus situations and objects. Just as a well-selected group of subjects must be representative of the population from which it was drawn, so a sample of stimulus situations must be representative of the total range of stimulus values in the environment of the organism. Experiments on size constancy again provide the paradigm. Instead of confining size judgments to a limited set of test stimuli at a few arbitrarily selected distances, Brunswik and his students have obtained representative samples of size judgments while their subjects went about the routine of their daily lives. A high degree of constancy in the appearance of physical bodies was shown to hold over a wide range of stimulus situations. At the same time, the importance of the attitude under which the judgments are made (whether a "betting" or a "critical" attitude) was demonstrated. This representative field study not only confirmed the results of laboratory experimentation but also served to give them a high degree of situational generality.

Situations as well as subjects, Brunswik affirms, must be subjected to representative sampling. If such a program were to be carried out, the experimenter's tools of statistical analysis must be adapted to the changing theory of design. The author suggests that correlational analysis may go a long way toward meeting this need.

The coefficient of correlation thus far has been largely a tool of differential psychology. A sample of subjects is exposed to two test situations or measurements and the covariation of these measurements is ascertained. The crucial point for Brunswik is that it is subjects who are sampled and that the range of situations over which the measurements are taken is highly restricted. As far as the logic of correlation is concerned, it is possible to sample situations and to correlate (1) different responses to the same situation, (2) responses and stimulus characteristics, and (3) different stimulus characteristics. To exemplify each of these, (1) we may correlate ratings of intelligence and ratings of personal appearance given by a single judge to a sample of individuals, (2) physical size (or retinal size) may be correlated with perceived size for a sample of objects, and (3) physical size may be correlated with retinal size. The important point is that such uses of the correlation coefficient are object-centered rather than subject-centered. As Brunswik himself puts it, "individuals and test situations have shifted places" (p. 34). Brunswik suggests that this use of the correlation coefficient may lead to a type of "intra-personal" factor analysis. In proposing that

individuals and test situations shift places, however, Brunswik does *not* mean the kind of shift involved in going from conventional factor analysis to inverse factor analysis. In conventional factor analysis, the correlation between two test variables or "features" is obtained over a sample of individuals with the situation held constant. What Brunswik suggests is that we obtain correlations between "features" for a single individual over a sample of situations. This type of analysis would, then, be an analysis of "intra-individual" correlations. Thus it may become possible to isolate the basic situational dimensions of perception by factorial methods. Any subject could be characterized in terms of situational factors which describe the dimensions of his perceptual world.

Object-centered correlational analysis leads Brunswik to a reexamination of the concepts of reliability and validity. He lays particular stress on indices of reliability which are based on correlations between responses: intra-individual and inter-individual reliability. Intra-individual reliability refers to the consistency of responses given by the same observer to the same sample of situations. It measures the stability of an individual's response to his environment. Inter-individual reliability—the agreement of observers in the presence of a given stimulus situation—measures the "ecological reliability of a response," i. e., the extent to which a stimulus has comparable psychological consequences for different organisms. Such measures of reliability again are object- rather than subject-centered. The concern is not with variability in a large sample of subjects but rather with the consistency of responses (given by one or two subjects) to an object or representative sample of objects.

Brunswik's conception of validity is rooted in his general theory of perception. Validity of a response refers primarily to what the response *achieves* for the organism. We must recall that Brunswik has analysed perceptual responses in terms of the "intentions" of the organism in responding to an object in the environment. Thus, in judging size, the observer may intend to attain the actual physical size of the object. The correlation between physical size and perceived size then gauges the extent to which this intention is realized. This correlation measures the "functional validity" of the judgment. Under a different attitude, the observer may intend to attain not the physical size but the retinal size of the object. In this latter case, the correlation between retinal size and judged size provides the measure of functional validity. This treatment of validity clearly identifies Brunswik as a modern functionalist. His fundamental concern as a behavior theorist is with the mechanisms mediating successful adjustment to the environment.

Perceptual experiment and theory provide Brunswik with the bulk of his illustrative material. He hopes, however, that the methodology for which he argues will be extended to other fields of inquiry, particularly learning. Learning experiments, he believes, no less than perceptual experiments, must meet the challenge of representativeness.

Brunswik aptly describes his approach to experimental design as "probabilistic functionalism." The watchword is representative sampling of objects and situations. "Psychology is conceived of as a fundamentally statistical discipline throughout its entire domain, with 'functional validity' taking its place alongside traditional test validity" (p.56). The generality and applicability of the findings will depend primarily on the representativeness of the samples.

Such, in broad outline, is Brunswik's thesis. The cogency of his argument hinges on the assumption that representativeness (ecological validity) is the touchstone of experimental design. Often systematic laboratory experiments fail,

indeed, to meet this criterion. Many such experiments are aimed at the construction of models and the isolation of basic mechanisms of behavior with the aid of these models. Representativeness then becomes largely a problem of application, and essentially parametric study. Underlying this philosophy of experimentation is the belief that simplification for the sake of analysis—the artificial restriction, isolation, and interlocking of variables—does not preclude the discovery of highly general, ecologically valid cause-effect relationships. Brunswik's own representative study of size constancy, for example, bore out in large measure the results of artificially restricted laboratory experiments.

Representative sampling of situations can be effective only if the universe of situations is sampled with respect to some crucial characteristic. It is in the discovery of such crucial characteristics that the systematic, classical type of experiment may well be propadeutic to Brunswik's type of representative design. It seems that representative surveys of behavioral situations cannot replace systematically controlled experiments; the two approaches must proceed side by side. Nor is the representative sampling of objects and situations necessarily and always superior to the representative sampling of subjects. It may often be important to vary the nature of the responding organism over a wide range, to sample different life histories and past experiences. In many respects, organisms are not interchangeable, and the sampling of situations with the nature of the organism constant or quasi-constant may fall short of representativeness.

We shall do well indeed to heed Brunswik's plea for representative sampling of behavioral situations. At the same time it must be hoped that the extension of statistical surveys of behavior would not weaken our concern with theoretical (albeit artificial) models.

Harvard University LEO POSTMAN